HeartFormer: Semantic-Aware Dual-Structure Transformers for 3D Four-Chamber Cardiac Point Cloud Reconstruction

Zhengda Ma[0009−0003−7057−4296] and Abhirup Banerjee[0000−0001−8198−5128]

Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, United Kingdom
zhengda.ma@eng.ox.ac.uk

Abstract. Accurate three-dimensional (3D) reconstruction of the four-chamber heart from routine cine MRI remains an open challenge due to sparse slice acquisition, inter-slice misalignment, and the structural heterogeneity of cardiac substructures. Although cine MRI is the clinical gold standard for cardiac assessment, its inherent 2D acquisition limits comprehensive 3D characterization of both ventricular and atrial anatomy. Existing reconstruction approaches predominantly focus on biventricular geometry and do not explicitly address multi-class anatomical completion under severe sparsity. We propose *HeartFormer*, a semantic-aware dual-structure transformer designed for multi-class cardiac point cloud completion. HeartFormer jointly models global cardiac topology and substructure-specific geometric priors within a unified coarse-to-fine refinement framework, enabling anatomically coherent four-chamber reconstruction from sparse and misaligned inputs. We embed HeartFormer within the first completely automated geometric deep learning pipeline for 3D four-chamber cardiac surface reconstruction from routine cine MRI. To facilitate systematic evaluation, we introduce *HeartCompv1*, the first large-scale public benchmark for multi-class cardiac point cloud completion, comprising 17,000 high-resolution 3D heart models. Extensive cross-domain experiments on HeartCompv1 and UK Biobank demonstrate that HeartFormer achieves state-of-the-art geometric accuracy while preserving clinically meaningful structural fidelity. Code and dataset are available at https://github.com/MultiMeDIA-Oxford/HeartFormer.

Keywords: 3D Cardiac Reconstruction · Cine MRI · Point Cloud Completion · Geometric Deep Learning · Multi-Class Anatomical Modeling.

1 Introduction

Cardiac magnetic resonance imaging (MRI) is widely regarded as the clinical gold standard for evaluating cardiac morphology and function [16, 21]. In routine practice, cine MRI acquires a limited set of two-dimensional (2D) short- and long-axis slices providing complementary multi-view observations of the heart; however, their inherently 2D slice-based nature constrains accurate 3D characterization of cardiac anatomy [4, 11]. High-fidelity 3D representations are essential for quantitative biomarker analysis, pathology visualization, and patient-specific simulations of cardiac mechanics [10, 19]. Existing reconstruction methods predominantly adopt a slice-based segmentation-to-surface paradigm, and remain susceptible to sparse spatial sampling and inter-slice motion [1, 14, 23, 27]. Recent advances in geometric deep learning approaches [3, 9, 26, 31, 35] largely focus on biventricular reconstruction as evidenced in the recent review [17], neglecting atrial anatomy whose thin walls and intricate topology pose additional challenges for accurate four-chamber reconstruction.

3D point clouds offer a flexible and efficient representation for modeling geometric structure, establishing a direct bridge between sparse sensor measurements and learning-based shape inference [13]. However, the inherent sparsity, incompleteness, and misalignment of acquired data make point cloud completion a fundamental problem in geometric deep learning [7, 30, 34]. Existing point cloud completion networks [20, 24, 28, 32] primarily address single-class objects or scene reconstruction, and do not explicitly account for structured multi-class completion within a single anatomical organ. In the context of cardiac reconstruction, this requires jointly modeling heterogeneous substructures and preserving their interdependent spatial relationships, an aspect that remains largely unexplored.

In this work, we address multi-class anatomical completion as a distinct and clinically motivated problem in geometric deep learning. Our objective is to simultaneously reconstruct six anatomically and functionally distinct cardiac substructures, namely left ventricular (LV) endocardium, LV epicardium, right ventricular (RV) endocardium, RV epicardium, left atrium (LA), and right atrium (RA), from sparse and misaligned cine MRI. This task entails substantial geometric heterogeneity within a single organ, including thin-walled atrial anatomy and tightly coupled inter-chamber spatial relationships, all of which must be recovered coherently to ensure anatomically valid four-chamber reconstruction.

To tackle these challenges, we introduce *HeartFormer*, a semantic-aware dual-structure transformer that jointly models global cardiac topology and substructure-specific geometric priors, enabling structurally consistent multi-class completion under severe sparsity and slice misalignment. HeartFormer is embedded within the first completely automated geometric deep learning pipeline for 3D four-chamber cardiac reconstruction from routine cine MRI, transforming multi-view 2D slices into sparse 3D point clouds and progressively refining them into dense, anatomically coherent surfaces. In addition, we release *HeartCompv1*, the first large-scale public benchmark for multi-class cardiac point cloud completion, comprising 17,000 high-resolution 3D heart models. This dataset establishes the first open resource for systematic evaluation of multi-class medical point cloud completion, facilitating fair comparison and large-scale supervised learning in this emerging area. Extensive cross-domain evaluation on HeartCompv1 and UK Biobank demonstrates HeartFormer's state-of-the-art (SOTA) geometric accuracy together with clinically meaningful structural fidelity.

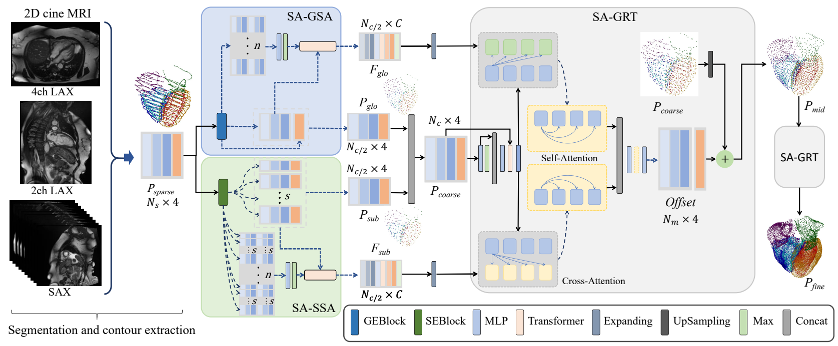

Fig. 1: Overview of the proposed HeartFormer architecture. 2D cine MRI slices are first segmented to extract cardiac contours, which are converted into sparse point clouds serving as the input. HeartFormer first employs a Semantic-Aware Global Structure Aggregation (SA-GSA) module and a Semantic-Aware Substructure Aggregation (SA-SSA) module to generate a coarse point cloud, $\mathbf{P}_{\text{coarse}}$, by modeling global context and substructure semantics. A subsequent two-stage SA-GRT progressively enhances geometric detail and anatomical consistency, yielding the final high-fidelity reconstruction, $\mathbf{P}_{\text{fine}}$.

2 Method

2.1 3D Four-Chamber Cardiac Reconstruction Pipeline

We present the first 3D four-chamber cardiac surface reconstruction pipeline based on point cloud completion, which converts misaligned 2D cine MRI slices into a complete 3D four-chamber cardiac point cloud and mesh. The pipeline takes a cine MRI stack comprising short-axis (SAX) and two- and four-chamber long-axis (2ch, 4ch LAX) views as input. First, a segmentation network generates the corresponding segmentation masks, from which contours are extracted [2]. These contours, combined with the spatial coordinates, are registered in 3D space and converted into point clouds, resulting in a sparse and misaligned cardiac point cloud. To handle sparse points across different substructures, we propose *HeartFormer*, a multi-class point cloud completion network that simultaneously corrects misalignments and infers dense, anatomically consistent point distributions. The completed point clouds enable high-quality mesh reconstruction through Poisson surface reconstruction followed by Ball Pivoting [5].

2.2 HeartFormer

The overall architecture of our proposed HeartFormer is illustrated in Fig. 1. We first employ semantic-aware dual-structure processing to generate an initial coarse point cloud, $\mathbf{P}_{\text{coarse}}$, that integrates both global and substructure-level information. Subsequently, a two-stage Semantic-Aware Geometry Refinement

Transformer (SA-GRT) serves as the core of the coarse-to-fine refinement module, leveraging global context and substructure semantics to progressively refine the point cloud. This process produces a dense, high-fidelity, and anatomically consistent cardiac point distribution, $\mathbf{P}_{\text{fine}}$.

Semantic-Aware Dual-Structure Processing. We introduce SA-GSA and SA-SSA modules to obtain a coarse labeled point cloud that integrates both global context and substructure-level information, while simultaneously extracting corresponding hierarchical features at the global and substructure levels to guide the subsequent refinement process, as illustrated in Fig. 1. The input sparse point cloud ($\mathbf{P}_{\text{sparse}} \in \mathbb{R}^{N_{\text{s}} \times 4}$) is first processed by SA-GSA to extract semantic-enhanced geometric representations, yielding a set of global key points $\mathbf{P}_{\text{glo}} \in \mathbb{R}^{N_{\text{c}/2} \times 4}$ and their corresponding global features $\mathbf{F}_{\text{glo}} \in \mathbb{R}^{N_{\text{c}/2} \times C}$. In parallel, the SA-SSA processes the input sparse point cloud to capture substructure-level key points $P_{sub} \in \mathbb{R}^{N_{\text{c}/2} \times 4}$ and substructure features $F_{sub} \in \mathbb{R}^{N_{\text{c}/2} \times C}$. N_{s} and N_{c} denote the numbers of sparse and coarse points, respectively, C is the feature dimension, and the 4 represents 3D coordinates with a semantic label.

$$(\mathbf{P}_{\text{glo}}, \mathbf{F}_1) = \mathcal{F}_{\text{GEBlock}}\big(\mathbf{P}_{\text{sparse}}\big), \tag{1}$$

$$\mathbf{F}_{\text{glo}} = \mathcal{T}\Big[\max_{\text{g}}\big(\mathcal{M}_{\text{g}}(\mathbf{F}_1)\big),\, \mathbf{P}_{\text{glo}}\Big], \tag{2}$$

$$(\mathbf{P}_{\text{sub}}, \mathbf{F}_2) = \mathcal{F}_{\text{SEBlock}}\big(\mathbf{P}_{\text{sparse}}\big), \tag{3}$$

$$\mathbf{F}_{\text{sub}} = \mathcal{T}\Big[\max_{\text{s}}\big(\mathcal{M}_{\text{s}}(\mathbf{F}_2)\big),\, \mathbf{P}_{\text{sub}}\Big], \tag{4}$$

where $\mathcal{F}_{\text{GEBlock}}$ and $\mathcal{F}_{\text{SEBlock}}$ denote the functions of GEBlock and SEBlock, respectively, which extract global and substructure geometric embeddings from the sparse input point cloud $\mathbf{P}_{\text{sparse}}$. The transformer operator $\mathcal{T}(\cdot)$ further aggregates contextual information to generate semantic-aware representations $\mathbf{F}_{\text{glo}}$ and $\mathbf{F}_{\text{sub}}$. The operators $\max(\cdot)$ and $\mathcal{M}(\cdot)$ represent the element-wise maximum aggregation and the multi-layer perceptron (MLP), respectively. The global and substructure representations are fused to form the initial coarse point cloud:

$$\mathbf{P}_{\text{coarse}} = [\mathbf{P}_{\text{glo}}, \mathbf{P}_{\text{sub}}], \tag{5}$$

where $[\,\cdot\,,\,\cdot\,]$ denotes feature-wise concatenation.

Semantic-Aware Geometry Refinement Transformer (SA-GRT). To further refine both the geometry and semantic distribution, we introduce a refinement strategy implemented via a two-stage SA-GRT. Learnable offset vectors are applied to the upsampled coarse points, enabling adaptive spatial adjustment and preserving semantic coherence during refinement. In the first stage, the coarse point cloud $\mathbf{P}_{\text{coarse}}$ is processed by a feature encoding module to obtain its latent representation:

$$\mathbf{F}_{\text{coarse}} = \Phi_{\text{E}}(\mathbf{P}_{\text{coarse}}), \tag{6}$$

where $\Phi_{\mathrm{E}}(\cdot)$ denotes the feature embedding function.

In parallel, the global features $\mathbf{F}_{\mathrm{glo}}$ and substructure features $\mathbf{F}_{\mathrm{sub}}$ are individually processed by their respective expanding functions. The expanded features, together with the coarse feature $\mathbf{F}_{\mathrm{coarse}}$, are fed into a Multi-Attention Aggregation (MAA) module that facilitates adaptive information exchange across structural levels.

$$\begin{aligned}\mathbf{F}'_{\mathrm{glo}} &= \mathcal{A}_{\mathrm{MAA}}\big(\Phi_{\mathrm{g}}(\mathbf{F}_{\mathrm{glo}}),\, \mathbf{F}_{\mathrm{coarse}}\big),\\ \mathbf{F}'_{\mathrm{sub}} &= \mathcal{A}_{\mathrm{MAA}}\big(\Phi_{\mathrm{s}}(\mathbf{F}_{\mathrm{sub}}),\, \mathbf{F}_{\mathrm{coarse}}\big),\\ \mathbf{F}'_{\mathrm{coarse}} &= [\,\mathbf{F}'_{\mathrm{glo}},\, \mathbf{F}'_{\mathrm{sub}}\,],\end{aligned} \tag{7}$$

where $\Phi_{\mathrm{g}}(\cdot)$ and $\Phi_{\mathrm{s}}(\cdot)$ denote the global and substructure expanding functions, respectively, and $\mathcal{A}_{\mathrm{MAA}}(\cdot)$ represents the MAA operator.

Given the enhanced feature representation $\mathbf{F}'_{\mathrm{coarse}}$, an offset generation module predicts fine-grained geometric adjustments, which are then applied to the upsampled coarse points to produce the intermediate point cloud $\mathbf{P}_{\mathrm{mid}}$:

$$\mathbf{P}_{\mathrm{mid}} = \mathcal{F}_{\mathrm{offset}}\big(\mathbf{F}'_{\mathrm{coarse}}\big) + \mathrm{Up}\big(\mathbf{P}_{\mathrm{coarse}}\big), \tag{8}$$

where $\mathcal{F}_{\mathrm{offset}}(\cdot)$ denotes the offset generation function and $\mathrm{Up}(\cdot)$ represents the upsampling operation.

Subsequently, the intermediate point cloud $\mathbf{P}_{\mathrm{mid}}$ is fed into the second stage of the SA-GRT, which performs another round of feature aggregation and offset refinement. This hierarchical refinement produces the final output $\mathbf{P}_{\mathrm{fine}}$:

$$\mathbf{P}_{\mathrm{fine}} = \mathcal{A}_{\mathrm{GRT}}(\mathbf{P}_{\mathrm{mid}}) + \mathrm{Up}\big(\mathbf{P}_{\mathrm{mid}}\big), \tag{9}$$

where $\mathcal{A}_{\mathrm{GRT}}(\cdot)$ denotes the feature aggregation operator within SA-GRT module.

Loss Function. For semantics-guided supervision, we employ a Semantic-Aware Chamfer Distance (SA-CD) loss that explicitly integrates anatomical semantics into geometric alignment. Unlike the standard Chamfer Distance, which treats all points equally, SA-CD computes the distance independently for each anatomical substructure, ensuring balanced optimization across components of varying sizes and point densities. By introducing semantic partitioning, this loss effectively preserves structural boundaries and mitigates cross-class interference during refinement. The SA-CD loss is defined as:

$$\mathcal{L}_{\mathrm{SA}}(\mathbf{P},\mathbf{G}) = \frac{1}{2K}\sum_{k=1}^{K}\Big(\mathbb{E}_{x\in\mathbf{P}^{(k)}}\min_{y\in\mathbf{G}^{(k)}}\|x-y\|_2 + \mathbb{E}_{y\in\mathbf{G}^{(k)}}\min_{x\in\mathbf{P}^{(k)}}\|x-y\|_2\Big), \tag{10}$$

where $\mathbf{P}^{(k)}, \mathbf{G}^{(k)} \in \mathbf{R}^{N\times 3}$ are the subsets of predicted and ground truth (GT) points corresponding to class k, $\mathbb{E}$ denotes the expectation operator. To enforce a coarse-to-fine reconstruction paradigm, SA-CD is applied at all prediction stages:

$$\mathcal{L} = L_{coarse} + L_{mid} + L_{fine} = \sum_{s\in\{\mathrm{coarse,\,mid,\,fine}\}} \mathcal{L}_{\mathrm{SA}}(\mathbf{P}_s,\mathbf{G}). \tag{11}$$

This hierarchical supervision jointly promotes global structural consistency and locally smooth surface reconstruction across all cardiac substructures.

2.3 Synthetic Dataset Generation

To address the scarcity of large-scale real-world 3D whole-heart anatomical datasets, we construct and publicly release *HeartCompv1* ($N = 17{,}000$), a large-scale dataset derived from the High-Resolution Atlas and Statistical Model [12]. Specifically, we sample full heart meshes from the statistical shape model by randomly varying the 50 principal modes of variation. We define slicing planes for each heart mesh in accordance with standard clinical cine MRI protocols [15]. Assuming slice misalignment can be modeled by rigid transformations, we introduce simulated misalignment artifacts caused by respiration or patient motion into each SAX and LAX slice. Following [4], misalignments are divided into five levels: none, mild, medium, strong, and severe. For each level, independent Gaussian perturbations with increasing standard deviations are applied to reflect varying degrees of misalignment. The resulting 2D contours are converted to 3D point clouds to generate sparse, misaligned contours representing real cine MRI. Finally, we extract dense point clouds from the vertices of the deformed heart meshes that serve as ground truth for training. We release the detailed dataset generation protocol and dataset on the GitHub repository and https://zenodo.org/records/20721166.

3 Experiments and Results

Dataset. HeartCompv1 dataset contains paired sparse and dense point clouds for training and evaluation. The training set consists of 10,000 samples with varying degrees of misalignment to improve model robustness. The validation set and five test sets each contain 1,000 samples. The five test sets correspond to different levels of misalignment, thereby representing different levels of reconstruction difficulty. In addition, we evaluate HeartFormer on a subset of the UK Biobank dataset [22] to assess its generalization in real-world scenarios. Specifically, we select a real-world clinical subset of 116 subjects, acquired from multiple imaging centers (Cheadle and Newcastle), to evaluate model performance under realistic heterogeneous clinical setting. The cohort includes both healthy and diseased individuals, as well as male and female participants (83/33), with a wide age range of 47–78.8 years (mean 65.5 ± 7.4 years). Multiple prevalent cardiovascular conditions are represented, including myocardial infarction (37.1%), hypertension (39.7%), ischemic heart disease (41.4%), atrial fibrillation (13.8%), and heart failure (6.0%).

Evaluation Metrics. To assess reconstruction accuracy, we employ Chamfer Distance (CD) and Hausdorff Distance (HD) as quantitative geometric metrics. In addition, we compute several clinically relevant indices, including standard volumetric measurements at end-diastole: left and right ventricular volumes (LVV, RVV) and left and right atrial volumes (LAV, RAV).

Evaluation on HeartCompv1 Dataset. We compare our proposed HeartFormer with SOTA single-class point cloud completion methods, including FoldingNet [30], PoinTr [33], FBNet [29], SnowflakeNet [25], AdaPoinTr [32], AnchorFormer [8], CRA-PCN [20], ODGNet [6], PointAttN [24], and SymmComp [28],

Table 1: Quantitative comparison of reconstruction results for five different levels of misalignment in the HeartCompv1 dataset. Best results are indicated in **bold**.

Methods	CD (mm) ↓					HD (mm) ↓				
	None	Mild	Medium	Strong	Severe	None	Mild	Medium	Strong	Severe
FoldingNet	2.174	2.233	2.350	2.489	2.674	16.323	16.457	16.573	16.660	16.996
SnowflakeNet	2.649	2.822	3.235	3.719	4.262	12.816	13.121	13.894	16.448	20.909
PoinTr	2.016	2.388	2.784	3.162	3.565	18.916	19.234	19.760	20.688	23.020
FBNet	1.453	1.710	1.939	2.147	2.369	7.738	8.446	10.141	11.788	13.722
AdaPoinTr	1.288	1.484	1.705	1.931	2.182	7.038	7.199	8.134	9.097	10.514
AnchorFormer	4.964	5.272	5.700	6.203	6.787	21.624	22.026	22.701	23.499	24.711
CRA-PCN	2.158	2.384	2.705	3.108	3.600	10.596	10.992	11.827	14.399	18.979
ODGNet	1.743	1.942	2.186	2.488	2.839	14.523	14.901	15.657	16.525	18.759
PointAttN	1.653	1.796	1.986	2.196	2.436	8.986	9.423	10.474	11.912	13.478
SymmComp	1.268	1.429	1.637	1.857	2.083	5.664	5.993	6.949	8.230	9.613
PCCN	1.628	1.746	1.992	2.323	2.731	6.723	7.065	8.277	9.761	11.662
HeartFormer	**1.194**	**1.373**	**1.585**	**1.799**	**2.037**	**5.417**	**5.678**	**6.285**	**7.401**	**8.662**

Table 2: Quantitative comparison of models on the UK Biobank dataset.

Structure	CD (mm) ↓		HD (mm) ↓	
	PCCN	Heart-Former	PCCN	Heart-Former
LV Endo	2.712	**1.382**	12.423	**11.184**
LV Epi	3.460	**1.551**	15.485	**12.057**
RV Endo	3.360	**1.790**	16.550	**15.924**
RV Epi	3.374	**1.755**	17.706	**11.456**
LA	5.036	**2.344**	20.398	**19.749**
RA	3.723	**2.037**	16.217	**15.038**

Table 3: 3D volumetric clinical metrics (mean ± SD) at end-diastole.

Sex	Metric	Proposed	Benchmark
Female	LVV	131.0 ± 11.4	124 ± 21
	RVV	142.3 ± 10.8	130 ± 24
	LAV	57.5 ± 10.4	62 ± 17
	RAV	62.8 ± 9.8	69 ± 17
Male	LVV	159.8 ± 19.3	166 ± 32
	RVV	185.2 ± 20.0	182 ± 36
	LAV	66.2 ± 12.3	71 ± 19
	RAV	92.8 ± 25.9	93 ± 27

as well as the multi-class completion method PCCN [4]. We evaluate all methods on five test sets exhibiting different degrees of misalignment: none, mild, medium, strong, severe. These configurations simulate realistic inter-slice motion artifacts commonly observed in dynamic cardiac MRI, enabling a systematic assessment of robustness of each model under varying conditions. For quantitative comparison in Table 1, our method achieves the best performance across all misalignment levels and metrics with statistically significant improvements (Wilcoxon signed-rank test). HeartFormer contains 5.688M parameters with 87.036 GFLOPs, and achieves an average inference time of 0.0358s per sample on a single GPU.

Analysis on UK Biobank Dataset. Table 2 further demonstrates that HeartFormer consistently and statistically significantly (Wilcoxon signed-rank test) outperforms PCCN for each cardiac substructure in real-world clinical scenarios. Table 3 summarizes the estimated end-diastolic 3D volumetric clinical metrics for female and male subjects, alongside the reference values from [18]. Our 3D reconstructions achieve physiologically plausible scores across all metrics and accurately capture sex-related differences. Figure 2 presents reconstruction results

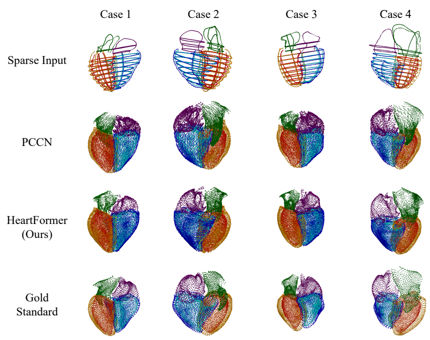

Fig. 2: Qualitative comparison on the UK Biobank dataset.

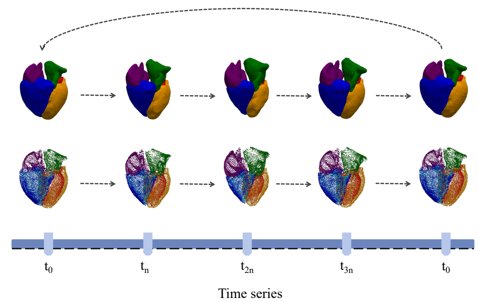

Fig. 3: 4D cardiac mesh and point cloud reconstruction for one UK Biobank subject. The dynamic video is in the supplementary.

Table 4: Effects of global and substructure geometric features and points.

F_{glo}	P_{glo}	F_{sub}	P_{sub}	CD ↓	HD ↓
✓	✓			1.701	7.411
✓	✓		✓	1.678	7.355
		✓	✓	1.689	7.390
	✓	✓	✓	1.648	7.267
✓	✓	✓	✓	**1.596**	**6.750**

Table 5: Effects of loss functions.

SA-CD					
L_{coarse}	L_{mid}	L_{fine}	L_{geo}	CD ↓	HD ↓
			✓	Worng labeling	
✓		✓		1.624	6.988
	✓	✓		1.637	7.071
✓	✓	✓		**1.596**	**6.750**
✓	✓	✓	✓	1.616	6.876

for four representative subjects from the UK Biobank dataset. The proposed HeartFormer produces more anatomically faithful and complete cardiac surfaces compared to PCCN. Figure 3 shows the 4D cardiac mesh and point cloud reconstruction for a typical UK Biobank subject, effectively capturing the dynamic motion of cardiac contraction and relaxation. Both qualitative and quantitative analyses confirm that HeartFormer delivers accurate, robust, and clinically meaningful reconstruction performance.

Ablation Studies. As presented in Table 4, both global and substructure geometric points and features substantially improve model performance. In particular, the substructure geometric points and features contribute more significantly than the global geometric components. In Table 5, removing any component in Semantic-Aware Geometry loss leads to a consistent drop in performance, confirming the necessity of multi-scale semantic alignment. Models trained without Semantic-Aware guidance show clear semantic mislabeling, while the additional geometry loss (i.e., the Chamfer Distance loss without Semantic-Aware guidance) does not yield further improvement, suggesting that the Semantic-Aware Geometry loss already provides strong geometric regularization.

We further analyzed the robustness of HeartFormer under defective and noisy input conditions. Specifically, we simulated incomplete point clouds with 10% and 20% missing points and introduced localized perturbations by adding Uniform noise within [-3, 3]mm to 10-20% of randomly selected points, mimicking segmentation inaccuracies commonly encountered in clinical practice. Compared with the standard setting (chamfer distance: 1.585 mm), HeartFormer maintained stable performance under missing-point conditions (10%: 1.588 mm; 20%: 1.599 mm) and noisy perturbations (10%: 1.611 mm; 20%: 1.646 mm), demonstrating its robustness and stability under challenging clinical-like conditions.

4 Conclusion

In this work, we present the first geometric deep learning framework for 3D four-chamber cardiac reconstruction from cine MRI using multi-class point cloud completion. Central to our framework is *HeartFormer*, a novel completion network that jointly leverages global anatomical context and local substructural priors, while adaptively refining substructure representations under semantic-conditioned guidance to achieve high-fidelity cardiac reconstruction. We introduce *HeartCompv1*, the first large-scale dataset for multi-class cardiac point cloud completion. We hope HeartFormer and HeartCompv1 will inspire future research in this direction and contribute to advancing clinically impactful 3D and 4D cardiac modeling.

Acknowledgments. This study was conducted using the UK Biobank resource under Application Number '135228'. The UK Biobank study protocol was approved by the North West Multi-Center Research Ethics Committee (MREC) as a Research Tissue Bank approval (REC reference: 11/NW/0382) and received approval extension on 2021 (REC reference: 21/NW/0157). The study complies with the Declaration of Helsinki, and participants signed a consent form prior to conducting the study. The authors acknowledge the use of the services/facilities of the Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford. The work of AB was supported by the Royal Society University Research Fellowship (grant no. URF\R1\221314) and the British Heart Foundation (BHF) Oxford Centre of Research Excellence (RE/24/130024).

Disclosure of Interests The authors have no competing interests to declare.

References

1. Banerjee, A., Camps, J., Zacur, E., 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 **379**(2212), 20200257 (2021)
2. Banerjee, A., Zacur, E., Choudhury, R.P., Grau, V.: Automated 3D whole-heart mesh reconstruction from 2D cine MR slices using statistical shape model. In: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). pp. 1702–1706. IEEE (2022)
3. Beetz, M., Banerjee, A., Grau, V.: Point2mesh-net: Combining point cloud and mesh-based deep learning for cardiac shape reconstruction. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 280–290. Springer (2022)

4. 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. Medical Image Analysis **90**, 102975 (2023)
5. 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)
6. Cai, P., Scott, D., Li, X., Wang, S.: Orthogonal dictionary guided shape completion network for point cloud. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 864–872 (2024)
7. Chen, Z., Wang, Y., Nan, L., Zhu, X.X.: Parametric point cloud completion for polygonal surface reconstruction. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 11749–11758 (2025)
8. Chen, Z., Long, F., Qiu, Z., et al.: Anchorformer: Point cloud completion from discriminative nodes. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 13581–13590 (2023)
9. Deng, Y., Xu, H., Rodrigo, S., Williams, S.E., Williams, M.C., Niederer, S.A., Pushparajah, K., Young, A.: Modusgraph: automated 3d and 4d mesh model reconstruction from cine cmr with improved accuracy and efficiency. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 173–183. Springer (2023)
10. Doste, R., Camps, J., Wang, Z.J., et al.: An automated computational pipeline for generating large-scale cohorts of patient-specific ventricular models in electromechanical in silico trials. Computer Methods and Programs in Biomedicine p. 109290 (2026)
11. Frangi, A.F., Niessen, W.J., Viergever, M.A.: Three-dimensional modeling for functional analysis of cardiac images, a review. IEEE Transactions on Medical Imaging **20**(1), 2–5 (2002)
12. Hoogendoorn, C., Duchateau, N., Sanchez-Quintana, D., et al.: A high-resolution atlas and statistical model of the human heart from multislice CT. IEEE Transactions on Medical Imaging **32**(1), 28–44 (2012)
13. Huang, Z., Wen, Y., Wang, Z., et al.: Surface reconstruction from point clouds: A survey and a benchmark. IEEE Transactions on Pattern Analysis and Machine Intelligence (2024)
14. Ma, Q., Meng, Q., Qiao, M., Matthews, P.M., O'Regan, D.P., Bai, W.: Cardiacflow: 3d+ t four-chamber cardiac shape completion and generation via flow matching. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 89–99. Springer (2025)
15. Margeta, J., Criminisi, A., Lee, D.C., Ayache, N.: Recognizing cardiac magnetic resonance acquisition planes. In: MIUA-Medical Image Understanding and Analysis Conference-2014 (2014)
16. Mavrogeni, S.I., Kallifatidis, A., Kourtidou, S., et al.: Cardiovascular magnetic resonance for the evaluation of patients with cardiovascular disease: An overview of current indications, limitations, and procedures. Hellenic Journal of Cardiology **70**, 53–64 (2023)
17. McMillian, E., Banerjee, A., Bueno-Orovio, A.: From 2D to 3D, deep learning-based shape reconstruction in magnetic resonance imaging: A review. arXiv preprint arXiv:2510.01296 (2025)
18. Petersen, S.E., Aung, N., Sanghvi, M.M., Zemrak, F., Fung, K., Paiva, J.M., Francis, J.M., Khanji, M.Y., Lukaschuk, E., Lee, A.M., 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 **19**(1), 18 (2017)
19. Rodero, C., Baptiste, T.M., Barrows, R.K., et al.: Advancing clinical translation of cardiac biomechanics models: a comprehensive review, applications and future pathways. Frontiers in Physics **11**, 1306210 (2023)
20. Rong, Y., Zhou, H., Yuan, L., et al.: CRA-PCN: Point cloud completion with intra-and inter-level cross-resolution transformers. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 4676–4685 (2024)
21. Scatteia, A., Dellegrottaglie, S.: Cardiac magnetic resonance in ischemic cardiomyopathy: present role and future directions. European Heart Journal Supplements **25**(Supplement_C), C58–C62 (2023)
22. UK Biobank: UK Biobank: protocol for a large-scale prospective epidemiological resource. Accessed May **7**(2016), 1–112 (2007)
23. Villard, B., Zacur, E., Grau, V.: ISACHI: integrated segmentation and alignment correction for heart images. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 171–180. Springer (2018)
24. Wang, J., Cui, Y., Guo, D., et al.: Pointattn: You only need attention for point cloud completion. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 5472–5480 (2024)
25. Xiang, P., Wen, X., Liu, Y.S., et al.: Snowflake point deconvolution for point cloud completion and generation with skip-transformer. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(5), 6320–6338 (2022)
26. Xiao, J., Zheng, W., Wang, W., et al.: Slice2Mesh: 3D surface reconstruction from sparse slices of images for the left ventricle. IEEE Transactions on Medical Imaging (2024)
27. Xu, H., Muffoletto, M., Niederer, S.A., Williams, S.E., Williams, M.C., Young, A.A.: Whole heart 3d shape reconstruction from sparse views: leveraging cardiac computed tomography for cardiovascular magnetic resonance. In: International Conference on Functional Imaging and Modeling of the Heart. pp. 255–264. Springer (2023)
28. Yan, H., Li, Z., Luo, K., et al.: SymmCompletion: High-fidelity and high-consistency point cloud completion with symmetry guidance. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 9094–9102 (2025)
29. Yan, X., Yan, H., Wang, J., et al.: FBNet: Feedback network for point cloud completion. In: European Conference on Computer Vision. pp. 676–693. Springer (2022)
30. Yang, Y., Feng, C., Shen, Y., Tian, D.: Foldinget: Interpretable unsupervised learning on 3d point clouds. arXiv preprint arXiv:1712.07262 **2**(3), 5 (2017)
31. Ye, M., Yang, D., Kanski, M., et al.: Neural deformable models for 3D bi-ventricular heart shape reconstruction and modeling from 2D sparse cardiac magnetic resonance imaging. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 14247–14256 (2023)
32. Yu, X., Rao, Y., Wang, Z., et al.: AdaPoinTr: Diverse point cloud completion with adaptive geometry-aware transformers (2023). arXiv preprint arXiv:2301.04545
33. Yu, X., Rao, Y., Wang, Z., et al.: PoinTr: Diverse point cloud completion with geometry-aware transformers. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 12498–12507 (2021)
34. Yuan, W., Khot, T., Held, D., et al.: PCN: Point completion network. In: 2018 international conference on 3D vision (3DV). pp. 728–737. IEEE (2018)
35. Zhang, C., Luo, Y., Wu, Y., et al.: Topology-preserving loss for accurate and anatomically consistent cardiac mesh reconstruction. arXiv preprint arXiv:2503.07874 (2025)