

IMPERIAL

Deep Learning Surrogates for Finite Element Modelling of Cardiac Biomechanics

Author:

Siyu Mu

Supervised by:

Dr Choon Hwai Yap

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Department of Bioengineering

Imperial College London

2026

Abstract

Cardiac biomechanics provides a quantitative framework for understanding cardiac function and enables patient-specific assessment beyond conventional imaging biomarkers. However, the clinical adoption of image-driven cardiac biomechanical modelling remains limited by the difficulty of reconstructing patient-specific cardiac motion from medical images and the high computational cost of Finite Element Analysis (FEA) for simulating cardiac mechanics throughout the cardiac cycle.

To address these challenges, this thesis develops a series of Machine Learning (ML)-based methods that accelerate key stages of the Left Ventricle (LV) biomechanical modelling pipeline.

First, *Elastix Fourier Deep Learning* (EFDL) is proposed as an unsupervised motion-tracking framework that combines image registration with Deep Learning (DL)-based temporal regularisation to recover cycle-consistent cardiac motion trajectories. The method enables reliable 2D+t and 3D+t motion tracking across multiple imaging modalities.

Building upon these motion priors, the thesis introduces *IMC-PINN-FE*, a motion-consistent Physics-Informed Neural Network (PINN) framework for volume-driven ventricular mechanics. By integrating governing-equation constraints with tracked motion information, the method enables efficient estimation of myocardial stiffness and active tension while reconstructing physiologically coherent Pressure-Volume (P-V) behaviour.

To overcome the scalability limitations of patient-specific optimisation, the thesis further develops graph-based surrogate models. *HeartUnloadNet* predicts unloaded LV geometries directly from end-diastolic configurations on heterogeneous tetrahedral meshes, substantially reducing the computational burden of inverse cardiac mechanics. Extending this concept, *CardioGraphFENet* (CGFENet) is proposed as a unified surrogate model for full-cycle LV mechanics. Using a volume- and time-driven mapping, the framework predicts displacement fields and chamber pressure while supporting both forward loading and inverse unloading.

Overall, this thesis establishes an integrated computational pipeline for motion tracking, physics-informed parameter estimation, unloaded-state recovery, and fast surrogate simulation, improving

the efficiency and scalability of image-driven cardiac biomechanics.

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Declaration of Originality

This thesis presents original ideas that were developed through collaboration between the author and their collaborators. The author affirms that, to the best of their knowledge, the work is original and that proper attribution has been provided wherever necessary. Contributions made by others have been acknowledged through co-authorship of the relevant papers described in this thesis.

Parts of this thesis have been disseminated through peer-reviewed publications and conference presentations, as listed in the *Publications and Conferences* section. Any material that has appeared previously has been appropriately revised and extended for inclusion in this thesis, and the original sources are cited accordingly.

The author is the primary writer of this document unless stated otherwise. All tables and figures were produced by the author, including the formatting and styling of the document, unless explicitly credited to other sources.

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at Imperial College London only. It has not been submitted, either in whole or in part, for any other degree or professional qualification.

The thesis formatting is adapted from an open-source L^AT_EX template by Giordano Scarciotti.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr Choon Hwai Yap, and my postdoctoral mentor, Dr Wei Xuan Chan, for their patient guidance, support, and encouragement throughout my PhD. Their advice, insight, and high standards have been invaluable to the development of this work.

To my wife, Xinyue (Zoey) Liu, you are the most important part of my PhD journey. I am deeply grateful for your unwavering support, patience, and understanding during the many demanding stages of this project. Your encouragement has sustained me through both challenges and milestones.

Finally, I would like to thank my family for their constant love and support. Their belief in me has been a continued source of motivation throughout this journey.

Contents

Abstract	i
Copyright	iii
Declaration of Originality	iv
Acknowledgments	v
List of Acronyms	x
List of Figures	xiii
List of Tables	xviii
1 Introduction	1
1.1 Background and Motivation	1
1.1.1 Image-based Cardiac Biomechanics and FEA	1
1.1.2 Computational Challenges in Personalised Cardiac FEA	2
1.1.3 DL for Image-driven Biomechanics	4
1.2 Research Aims	5
1.3 Thesis Outline	6
1.4 Publications and Conferences	6
2 Related Work	8
2.1 Introduction	8
2.2 Cardiac Image Registration and Motion Tracking	9
2.2.1 2D Tracking Techniques	9
2.2.2 3D Tracking Techniques	12
2.3 Finite Element Analysis of Left Ventricular Mechanics	14
2.3.1 Principles	14
2.3.2 Clinical use of FEA	17
2.4 ML Surrogates in Cardiac Biomechanics	19
2.4.1 Early ML Surrogates	19

2.4.2	PINN-based Surrogates	20
2.4.3	Graph-based Surrogates	25
3	Unsupervised Cross-Modal Motion Tracking over the Cardiac Cycle	28
3.1	Introduction	29
3.1.1	Contributions	31
3.2	Methods	31
3.2.1	B-spline Pairwise Registration	32
3.2.2	Loss function design: Fourier regularization	33
3.2.3	Connection Methods	34
3.3	Experiments and Results	35
3.3.1	Datasets	35
3.3.2	Evaluation factors	37
3.4	Experiments and Results	37
3.4.1	2D Evaluation	37
3.4.2	3D Evaluation	42
3.5	Discussion	42
3.6	Conclusion	44
4	Reduced-Order Physics-Informed Neural Networks for Left Ventricular Mechanics over the Cardiac Cycle	45
4.1	Introduction	46
4.1.1	Contributions	48
4.2	Methods	49
4.2.1	Myocardial Stiffness and Maximal Tension Estimator	50
4.2.2	Volume-Constrained PINN-FEA Solver	55
4.2.3	Neural Network Architecture	57
4.3	Experiments and Results	60
4.3.1	Performance of $C_{\text{stiffness}}$ and T_{max} Estimator	60
4.3.2	Performance of PINN-FEA solver	62
4.4	Discussion	67
4.4.1	Sensitivity and Error Analysis	68
4.4.2	Limitations and Future Work	69
4.5	Conclusion	70

5	Graph Neural Networks for Left Ventricular Unloading	72
5.1	Introduction	73
5.1.1	Contributions	75
5.2	Methods	75
5.2.1	Finite Element Ground Truth Generation	75
5.2.2	HeartUnloadNet Architecture	76
5.2.3	Experimental Setup	81
5.3	Experiments and Results	85
5.3.1	Quantitative Performance Comparison	85
5.3.2	Reduced Supervised Learning Performance	86
5.3.3	Ablation Studies	87
5.3.4	Generalization to Unseen Global Physiological Values	89
5.3.5	Qualitative Visualization of Predictions	90
5.3.6	Advantages over Principal Component Analysis (PCA)-Based and Inverse FEA Approaches	91
5.4	Discussion	92
5.4.1	Novelty and Contribution	92
5.4.2	Implications	93
5.4.3	Architectural Strengths and Outlook	93
5.4.4	Limitations and Future Work	94
5.5	Conclusion	96
6	Graph Neural Network Surrogates for Left Ventricular Mechanics over the Car- diac Cycle	97
6.1	Introduction	98
6.1.1	Contributions	99
6.2	Methods	100
6.2.1	Problem Domain	100
6.2.2	CardioGraphFENet	102
6.2.3	Lumped parameter model coupling and Pressure-Volume-Time (P-V-t) lookup table	110
6.2.4	Data Generation	111
6.2.5	Experimental Setup	113
6.3	Results	113
6.3.1	Overview	113

6.3.2	Baseline Comparison	114
6.3.3	Ablation Study	116
6.3.4	Reduced supervision capability	119
6.3.5	Windkessel-coupled P-V loop evaluation	120
6.4	Discussion	121
6.4.1	Overall performance and physiological fidelity	121
6.4.2	Comparison with existing cardiac surrogate literature	123
6.4.3	Limitations and future work	124
6.5	Conclusion	125
7	Conclusion	126
7.1	Summary of contributions	126
7.2	Limitations and implications	127
7.3	Future directions	128
7.4	Closing remarks	128
8	Appendix	129
8.1	Helix-Angle Definition for IMC-PINN-FE II (Biomechanical Parameter Estimation)	129
8.2	Default FEA Parameters in <i>IMC-PINN-FE</i>	130
8.3	Weights of Loss Components in <i>IMC-PINN-FE</i> III	130
8.4	Evaluated factors in <i>IMC-PINN-FE</i>	130
8.5	Sensitivity and Error Analysis of <i>IMC-PINN-FE</i>	132
8.5.1	Number of Temporal Frames	132
8.5.2	Robustness of the 1/3 Diastolic Unloading Reference	133
8.5.3	Cavity Volume Perturbation	135
8.5.4	Sensitivity of Imaging Noise	136
8.5.5	Loss Component Comparison	138
8.5.6	Reintroducing Isovolumetric Phases via a Windkessel-Constrained Volume– Time Curve	138
8.6	Data Used for Pre-trained Surrogate Comparison in <i>IMC-PINN-FE</i>	140

List of Acronyms

AD Automatic Differentiation

BC Boundary Condition

BSF B-spline Fourier

CAP Cardiac Atlas Project

CI Confidence Interval

CMR Cardiovascular Magnetic Resonance

CNN Convolutional Neural Network

CRT Cardiac Resynchronization Therapy

CT Computed Tomography

DL Deep Learning

DLIR Deep Learning Image Registration

DSC Dice Similarity Coefficient

DT Diffusion Tensor

ECG Electrocardiogram

ED End-Diastole

EDP End-Diastolic Pressure

EDV End-Diastolic Volume

EF Ejection Fraction

ECHO Echocardiography

ES End-Systole

ESV End-Systolic Volume

FCNN Fully Connected Neural Network

FEA Finite Element Analysis

GAT Graph Attention Network

GCN Graph Convolutional Network

GNN Graph Neural Network

GP Gaussian Process

GRU Gated Recurrent Unit

HD Hausdorff Distance

HPC High-Performance Computing

IC Initial Condition

LOVO Leave-One-Value-Out

LV Left Ventricle

LVEDP Left Ventricular End-Diastolic Pressure

LVH Left Ventricular Hypertrophy

MC Monte Carlo

MD Mean Distance

ML Machine Learning

MLP Multi Layer Perceptron

MRI Magnetic Resonance Imaging

MSE Mean Square Error

NNFE Neural Network Finite Element

OBS Observation

PCA Principal Component Analysis

PDE Partial Differential Equation

PINN Physics-Informed Neural Network

POD Proper Orthogonal Decomposition

PS Peak-Systolic

PSP Peak-Systolic Pressure

P-t Pressure-Time

P-V Pressure-Volume

P-V-t Pressure-Volume-Time

RE Relative Ratio

RMSE Root Mean Squared Error

RNN Recurrent Neural Network

RV Right Ventricle

SD Standard Deviation

SOTA State-of-the-Art

SR Supervision Ratio

STIC Spatio-Temporal Image Correlation

SVD Singular Value Decomposition

V-t Volume-Time

List of Figures

2.1	Representative biomechanical outputs from patient-specific FEA in fetal cardiac intervention. The top row shows LV P-V loops, and the bottom row shows the corresponding myofibre stress trajectories over the cardiac cycle, for three representative fetal cases: BV2, UV2, and H3. BV2 and UV2 denote cases with later biventricular and univentricular outcomes, respectively, while H3 denotes a healthy control. The BV and UV cases exhibit distinct P-V behaviour and markedly different peak systolic myofibre stress, with the BV case showing a substantially higher stress peak than the UV case. These differences are consistent with the study finding that pre-intervention peak systolic myofibre stress is a strong predictor of fetal aortic valvuloplasty outcome [57].	18
2.2	FEA comparison of candidate pacing sites for Cardiac Resynchronization Therapy (CRT). Panel (a) shows the predicted changes in left ventricular pressure (ΔP_{LV}) and activation time (Δt), while panel (b) shows the predicted changes in the maximum rate of pressure rise ($\Delta dP/dt$) and left ventricular ejection fraction (ΔEF). Each coloured bar represents one candidate pacing location evaluated in the model [72].	19
2.3	Vanilla PINN schematic in cardiac mechanics. Spatial coordinates and problem-related parameters are provided as inputs to a Fully Connected Neural Network (FCNN), which predicts displacement fields throughout the computational domain. Automatic Differentiation (AD) is then used to compute spatial derivatives required by the governing equations. The network is trained by minimising a composite loss function consisting of physics-based residuals together with boundary conditions, initial conditions, and observational constraints.	21
2.4	Two representative Graph-based surrogates for LV mechanics: [18] employed an encoder–Graph Neural Network (GNN) processor–decoder architecture to model LV deformation, [19] extended this framework to a PI-GNN architecture, incorporating biomechanical constraints during training.	26
3.1	Schematic of the <i>EFDL</i> Framework: The input consists of 3D/4D full-cycle cardiac images (Echocardiography (ECHO)/Magnetic Resonance Imaging (MRI)). Initially, these images are processed through a rapid elastix pairwise registration module, which generates B-spline coefficients delineating the approximate transformations between each pair of neighbouring images. Subsequently, Neural Network 1, functioning as an encoder, reduces the dimensionality of the features. Neural Network 2, the primary network, executes Fourier regularization. The output, comprising Fourier coefficients, facilitates the back-computation of displacements across all time points. This enables the tracking of cyclic movements in the myocardium, with the resultant tracking outcomes displayed.	30
3.2	Connection approaches: MIX, ADD, and MODE.	34
3.3	Representative tracking visualisations (left) and per-frame Dice Similarity Coefficient (DSC) trajectories (right) for Adult ECHO, Fetal ECHO, and Adult MRI over one cardiac cycle. The reference frame used for deformation has DSC = 1 in each plot.	41

4.1	The overall framework of <i>IMC-PINN-FE</i>. Inputs to the framework include 4D (3D+time) imaging data (e.g., ECHO or MRI) combined with LV reconstruction obtained via DL and physiological inputs, including Peak-Systolic (PS) and End-Diastole (ED) LV pressures. Outputs include estimates of patient-specific LV myocardial stiffness and maximum active tension, and LV FEA simulation results across the cardiac cycle.	49
4.2	Workflow of the 3D <i>EFDL</i> in <i>IMC-PINN-FE</i>: The input consists of 4D full-cycle cardiac images (ECHO/MRI). Initially, these images undergo processing through a rapid Elastix pairwise registration module, which computes B-spline coefficients representing the approximate transformations between consecutive image pairs for coarse registration. Next, a neural network applies Fourier regularization using coordinate-based inputs. The output, comprising Fourier coefficients, enables the reconstruction of displacements across all time points. When combined with manual segmentation of the initial frame t_0 using LazySnap [107] for coarse delineation and Geomagic (https://oqton.com/geomagic-designx) for 3D refinement, this framework enables the tracking of cyclic myocardial motion.	50
4.3	Schematic of validation case generation. A limited number of states (blue points) are selected from the FEA simulation, which is performed using a given unloaded surface mesh (red points), to serve as tracking results.	55
4.4	Workflow of the <i>IMC-PINN-FE</i> framework. I. Deep Learning Motion Tracking: 4D full-cycle cardiac images (Echo/MRI) are used to generate myocardial motion meshes via either <i>EFDL</i> (our proposed method, no pre-training required) or Co-attention STN (previous SOTA method, pre-training required). II. Myocardial Stiffness and Maximal Tension Estimator: Starting from the 3D+t surface mesh, the pipeline includes: (1) tetrahedral meshing via Gmsh; (2) preprocessing to define helix angles and the load-free reference state; (3) Proper Orthogonal Decomposition (POD) to extract motion bases and amplitudes across frames; and (4) personalized estimation of myocardial stiffness ($C_{\text{stiffness}}$) and maximal active tension (T_{max}) using ED and PS amplitudes. III. Volume-Constrained PINN-FEA Solver: Inputs include myocardial volumes and their time points (from Section 2.1), sampled uniformly from the cardiac cycle starting at ED. A PINN with MLP architecture predicts LV pressure P_t and motion amplitudes $A_{n,t}$, from which displacement fields u_t are reconstructed via Eq. 4.2, enabling personalized P-V loop and biomechanical analysis.	59
4.5	P-V loop evaluation for (1) adult-2 and (2) fetal-2 cases, comparing <i>IMC-PINN-FEA</i> , FEA simulation, fixed image tracking, and Buoso et al.'s PINN-FEA [22]. Blue represents the P-V loop predicted by <i>IMC-PINN-FE</i> , orange represents the P-V loop from FEA simulation, gray represents the P-V loop generated by fixed displacements from image tracking, and red represents the P-V loop predicted by Buoso et al.'s PINN-FEA [22].	63
4.6	Box plots of DSC evaluation (20 uniformly sampled frames across the cardiac cycle for each of the 6 cases), comparing LV geometries from <i>IMC-PINN-FE</i> , traditional FEA, and Buoso et al.'s PINN-FEA [22] against image-derived results. Orange indicates the DSC between FEA outputs and image tracking, blue indicates <i>IMC-PINN-FE</i> vs. image tracking, and red indicates Buoso et al.'s PINN-FEA vs. image tracking.	65
4.7	Pointwise deformation error maps across different time points throughout the cardiac cycle (Adult-3). Each row corresponds to a different method: <i>IMC-PINN-FE</i> , traditional FEA, and Buoso et al.'s PINN-FEA [22]. Color encodes the magnitude of pointwise error with respect to the image-derived motion (contour). .	66

5.1	Schematic of the proposed <i>HeartUnloadNet</i> for estimating the unloaded LV configuration from an ED state. Inputs comprise the ED myocardial tetrahedral mesh (“Mesh parameters”) and global physiological parameters (“Global parameters”, including ED pressure, Fung stiffness scaling coefficient and helix angles). A shared graph attention-based encoder processes the mesh, while a global-parameter encoder embeds the scalar inputs, their features are fused by cross-attention and layer normalisation, yielding a node-wise representation conditioned on both local geometry and global physiology. On the unloading workflow, a decoder maps the fused features to nodal displacements and the corresponding unloaded myocardium, which are constrained by a supervised loss $L_{\text{supervised}}$ against the FEA inverse-unloading reference. On the loading workflow, the unloaded configuration is passed through the same encoder and a loading decoder to reconstruct the ED myocardium, with a cycle-consistency loss L_{cycle} enforcing agreement with the ED ground truth. The network is trained by minimising $L_{\text{total}} = \lambda_{\text{cycle}}L_{\text{cycle}} + L_{\text{supervised}}$. Colours indicate functional roles: green for inputs/outputs, blue for encoders and feature fusion, orange and purple for unloading and loading decoders, red for the cycle-consistency path, and black boxes for loss terms.	77
5.2	Architecture of the graph attention-based encoder used in <i>HeartUnloadNet</i>. The input is the ED myocardial mesh represented as a graph $\mathcal{M}_{\text{ED}} = (\mathbf{X}_e, \mathcal{E}_e)$, where $\mathbf{X}_e$ are node features and $\mathcal{E}_e$ the edge set. Node features are successively updated by three GATConv layers with residual connections (“ $\oplus$ ”). A global mean pooling operation then aggregates the node-wise features into a graph-level descriptor, which is combined with the node features in a MultiheadAttention block using query (Q), key (K) and value (V) projections. The output of this attention module, with an additional residual connection and layer normalisation, provides the fused latent representation of the ED myocardium used by the downstream decoders.	79
5.3	Qualitative heatmap visualization of vertex-wise prediction errors on two representative test cases. Each row shows a different method (Inverse FEA, PointNet++, Graph Convolutional Network (GCN) and <i>HeartUnloadNet</i>), and each column corresponds to one test case. Warmer colors indicate higher local displacement errors.	91
5.4	Qualitative heatmap visualization comparing vertex-wise prediction errors of PCA and <i>HeartUnloadNet</i> for a representative test case. Compared to the ground truth, PCA exhibited a Hausdorff Distance (HD) of 1.02 ± 0.07 cm, whereas <i>HeartUnloadNet</i> achieved 0.083 ± 0.028 cm.	92
6.1	Overview of the problem domain and the role of <i>CGFENet</i>: forward loading maps prescribed (V, t) and unloaded geometry Ω_0 to pressure and displacement. Inverse unloading recovers Ω_0 from a loaded geometry. <i>CGFENet</i> replaces both expensive FEA operators with a unified cycle-consistent surrogate.	100
6.2	Overall architecture of <i>CGFENet</i>: A dual-stream design encodes LV geometry (graph encoder) and the Volume-Time (V-t) signal (temporal encoder) into a shared latent space. The fused latent drives two task heads for forward loading (pressure and displacement) and inverse unloading (zero-pressure recovery), while a cycle-consistent constraint regularises the two mappings across the cardiac cycle. A pre-trained mesh-parameter estimator (frozen) provides the shared global features used in Fig. 6.3.	102

6.3	Schematic of the Graph Fusion Encoder: Node features include both node-specific attributes and shared global descriptors, and are processed through stacked GATv2 residual blocks. A global graph descriptor is obtained by mean pooling and fused back into node embeddings via cross-attention, enabling the encoder to integrate local mesh connectivity with chamber level mechanical context.	103
6.4	Schematic of the Temporal Recurrent Neural Network (RNN) Encoder: The input sequence encodes the target chamber volume, its deviation from the reference volume, and a Fourier time representation. An initial Multi Layer Perceptron (MLP) generates time-conditioned latent vectors, which are propagated through a Gated Recurrent Unit (GRU) to produce the final temporal latent used for latent fusion.	106
6.5	Pretrained MLP-based mesh parameter estimator used to compute global LV descriptors.	107
6.6	Schematic of the lumped parameter model employed in the simulations. <i>C</i> : capacitance (compliance); <i>R</i> : resistance.	110
6.7	Overview of the data-generation pipeline: (a) PCA-based sampling of ED LV shapes, (b) construction of subject-specific tetrahedral ED meshes, (c) inverse-FEA computation yielding the unloaded, zero-pressure configuration with assigned fibre architecture, and (d) state-space FEA simulations producing displacement and pressure fields across prescribed cavity volumes.	111
6.8	(a) Distribution of ED LV cavity volumes and (b) myocardial wall thickness metrics (mean, minimum, maximum) across all synthesized meshes. The three wall-thickness distributions represent statistics computed over the myocardial wall surface of the same heart. These distributions are included to demonstrate the physiological plausibility and reliability of the generated dataset, with the wall-thickness values remaining within ranges reported in prior studies. The red dashed lines denote reference ranges derived from the literature, including normal adult ED LV volumes (approximately 110–190 mL) [164] and ED myocardial wall thickness spanning roughly 3–13 mm [165].	112
6.9	Pressure-Time (P-t) trajectories for a representative test case under several prescribed target volumes, comparing the ground-truth FEA solution (GT) with predictions from <i>CGFENet</i>, <i>MeshGraphNet</i>, and <i>GraphUNet</i>. In each panel, <i>CGFENet</i> most closely follows the target curve and yields the smallest volume errors, while the baseline surrogates exhibit larger deviations from the ground truth.	116
6.10	Reduced-supervision study: Vertex-level accuracy versus training epoch under Supervision Ratio (SR) of 70%, 50%, 30%, and 10%. <i>Cycle</i> uses all training cases: labelled samples contribute supervised and cycle-consistency losses, while unlabelled samples contribute only through cycle consistency. <i>No cycle</i> also processes all cases but updates the network using labelled samples only (unlabelled samples provide no loss/gradients).	120
6.11	Windkessel-coupled full-cycle P-V loops for four representative cases: <i>CGFENet</i> predictions are compared against FEA references after iterative surrogate-Windkessel coupling until a stable closed loop is reached.	122
6.12	Qualitative comparisons of ED and End-Systole (ES) deformation fields for four representative cases. Colour maps (<i>CGFENet</i> and FEA) show the deformation magnitude relative to the unloaded reference configuration, while <i>Diff</i> reports the residual error (prediction minus FEA).	123

-
- 8.1 **Sensitivity Analysis of sampling intervals (Adult-3):** 1) P-V loops at different number of frames 2) Estimated C (red) and T_{max} (blue) with error percentages . . . 132
- 8.2 **Geometric comparison of the unloaded state between the one third diastolic assumption (blue surface) and that obtained from inverse FEA unloading (red surface),** for (a) a representative healthy adult LV, and (b) a representative Left Ventricular Hypertrophy (LVH) pig case. A good agreement is found although a visible mismatch is at the apical region is found for (b). (c) Relative errors in cavity volume, myocardial volume, mean wall thickness and DSC between the unloaded LV from the two approaches, for 4 hearts (human adult, human fetal, healthy pig and LVH pig models). Small deviations were observed for most cases but there were moderate discrepancies in the LVH group. 133
- 8.3 **P-V loops under unloaded volume perturbations (healthy adult case):** The blue solid line shows the 0% baseline, and the grey dashed line shows the inverse FE reference. The green, yellow, orange and dark red curves correspond to 3%, 5%, 10% and 15% increases in unloaded cavity volume, respectively, illustrating that small perturbations (3%–5%) have little effect on the loop shape, whereas larger perturbations (10%–15%) mainly shift the loop near end diastole and alter end diastolic pressure. 134
- 8.4 **Sensitivity of POD motion modes to imaging noise level.** Low-order modes (1–4) retain DSC values close to 1 and show high robustness even at moderate noise levels, whereas higher-order modes (5–6) are more sensitive and exhibit noticeable degradation only at the extreme 50% noise level. 137
- 8.5 **Loss comparison across adult (1–3) and fetal (4–6) cases.** Subplots 1&4 show the total energy loss ($\text{Loss}_{\text{energyT}}$), subplots 2&5 show the relative pressure loss at peak systole ($\sqrt{\text{Loss}_{\text{PS}}}/P_{\text{GT}}$), and subplots 3&6 show the average relative volume loss ($\sqrt{\text{Loss}_{\text{volume}}}/V_{\text{GT}}$). *IMC-PINN-FE* consistently achieves low physics losses ($\text{Loss}_{\text{energyT}}$ and Loss_{PS}), comparable to the FEA baseline and substantially lower than Fixed Tracking. 139
- 8.6 **Recovering isovolumetric phases via a Windkessel-constrained volume–time curve.** Left: chamber volume–time trajectory for a representative healthy adult LV obtained from an FEA–Windkessel closed-loop simulation, showing a smooth profile with short near-constant-volume segments around end-systole and end-diastole that correspond to isovolumetric contraction and relaxation. Right: comparison between the reference FEA–Windkessel P-V loop and the loop reconstructed by *IMC-PINN-FE* when driven by the same Windkessel-derived V-t curve, illustrating close agreement in overall loop shape, End-Diastolic Pressure (EDP) and Peak-Systolic Pressure (PSP), while clearly preserving the isovolumetric segments. 139

List of Tables

3.1	Ablation study on N, B-grid size, and connection method. Row 1 is the reference setting.	38
3.2	Quantitative motion tracking results under single-modal, cross-dataset, and mixed-dataset training.	39
3.3	Quantitative comparison of 3D+t cardiac motion estimation methods on Adult MRI and Fetal ECHO. Time is training time	43
4.1	Physiological parameters of the 2 clinical image-tracking cases and 8 FEA-simulated validation cases. End-Diastolic Volume (EDV) and End-Systolic Volume (ESV) are derived from 3D image reconstructions. EDP and PSP for image tracking cases follow physiological values reported in prior clinical and experimental studies [130–132].	60
4.2	Myocardial stiffness ($C_{\text{stiffness}}$) estimation across 2 clinical image-tracking cases and 8 FEA-simulated validation cases. MC Dropout with 50 stochastic forward passes is used to compute a 95% Confidence Interval (CI). Relative Ratio (RE) denotes relative error relative to reference values.	60
4.3	Active tension (T_{max}) estimation across 2 clinical image-tracking cases and 8 FE-simulated validation cases. Monte Carlo (MC) Dropout with 50 stochastic forward passes provides a 95% CI. RE denotes relative error relative to reference values.	61
4.4	Comparison of State-of-the-Art (SOTA) surrogate models evaluated on the adult-only test set.	65
5.1	Simulation parameters used in the FEA pipeline. Pressure, fiber orientation, and the Fung stiffness scaling coefficient C were systematically varied, while the remaining dimensionless stiffness ratios b_{ff}, b_{xx}, b_{fx} were fixed across all simulations.	83
5.2	Evaluation metrics used for geometric and topological comparison.	85
5.3	Quantitative performance comparison of baseline methods for predicting unloaded LV geometry from the ED state.	85
5.4	Performance under varying levels of supervision. We compared the full model <i>HeartUnloadNet</i> (A0) with a variant (A1) that removes the cycle consistency loss. Each model was trained under decreasing SR, from full (100%) to extreme low-supervision (1%).	87
5.5	Stepwise ablation study of key architectural components in <i>HeartUnloadNet</i>. Models A1–A5 were derived by progressively removing individual modules from the full model A0, highlighting the contribution of each component to overall performance.	88
5.6	Component-wise Ablation Study of key architectural components in <i>HeartUnloadNet</i>. Models C0–C6 replaced encoder, decoder, or pooling modules.	88

5.7	Leave-One-Value-Out (LOVO) cross-validation results. Each variable group (pressure, stiffness, fiber orientation) is excluded during training to evaluate generalization.	90
6.1	Key training configurations.	114
6.2	Key FE constitutive parameters. The parameter settings are adapted from the literature [70] with minor adjustments.	114
6.3	Evaluation metrics used for geometric and pressure-based assessment.	115
6.4	Comparison with representative graph-based surrogate models on forward loading prediction. FEA served as ground truth.	116
6.5	Comparison of model size and inference runtime for generating a full case-level P-V-T map.	117
6.6	Displacement Ablation	118
6.7	Pressure Ablation	119
6.8	Unloaded state LV cavity volume and myocardial wall-thickness descriptors for four representative cases, including cavity volume V_V and wall thickness statistics (W_{mean} , W_{min} , W_{max}) computed from endocardial-epicardial distances.	121
6.9	Global LV functional indices derived from the Windkessel-coupled P-V loops for four representative cases. EDV: end-diastolic volume, ESV: end-systolic volume, Ejection Fraction (EF): ejection fraction. Δ denotes the absolute difference between <i>CGFNet</i> and FEA, with ΔEF reported in percentage points (pp).	121
6.10	Comparison across representative cardiac surrogate models at the data level	125
6.11	Comparison across representative cardiac surrogate models at the model level	125
8.1	Experimental endocardial and epicardial prescribed helix angle settings in adult, fetal, and pig models.	130
8.2	Default FEA parameters for adult, fetal, and pig models.	131
8.3	Weights of components for Volume-constrained PINN-FEA solver	131
8.4	Quantitative comparison of parameter back computation and P-V loop estimation when using the one third diastolic configuration as the unloaded state, with the inverse FEA unloading used as reference. C-RE: relative error of myocardial stiffness, T_{max} -RE: maximal active tension, P-V R^2 : coefficient of determination for the P-V relation, EDP-RE: relative errors in end diastolic pressure, PSP-RE: peak systolic pressure.	133
8.5	Sensitivity of parameter back-computation and P-V estimation to errors in the assumed unloaded-state cavity volume, compared to the 0% unloaded-volume baseline. C-RE: relative error of myocardial stiffness; T_{max} -RE: maximal active tension; P-V R^2 : coefficient of determination for the P-V relation; EDP-RE: relative error of end-diastolic pressure; PSP-RE: relative error of peak systolic pressure.	136

8.6	Sensitivity of <i>IMC-PINN-FE</i> parameter estimation and P-V estimation to imaging noise. Relative errors in myocardial stiffness (C-RE) and maximal active tension ($T_{\max}$ -RE), the coefficient of determination for the P-V relation (P-V R^2), and relative errors in end diastolic pressure (EDP-RE) and peak systolic pressure (PSP-RE) are reported for different noise levels, computed with respect to the noise-free reconstruction.	137
-----	---	-----

1

Introduction

Contents

1.1 Background and Motivation	1
1.1.1 Image-based Cardiac Biomechanics and FEA	1
1.1.2 Computational Challenges in Personalised Cardiac FEA	2
1.1.3 DL for Image-driven Biomechanics	4
1.2 Research Aims	5
1.3 Thesis Outline	6
1.4 Publications and Conferences	6

1.1 Background and Motivation

1.1.1 Image-based Cardiac Biomechanics and FEA

Cardiovascular disease remains one of the leading causes of morbidity and mortality worldwide, motivating continued efforts to improve mechanistic understanding and clinical management of cardiac dysfunction. While medical imaging modalities such as MRI and ECHO provide detailed anatomical and functional information, many determinants of cardiac performance, including myocardial stress and load redistribution, are not directly accessible through non-invasive measurements [1]. Characterising these mechanical quantities is therefore important for interpreting ventricular function and tissue-level disease progression, as well as for supporting clinical decision-making. For example, regional stress distributions can help identify mechanically vulnerable myocardium, evaluate the progression of ventricular remodelling, predict response to interventions such as cardiac

resynchronization therapy or valve repair, and support patient-specific treatment planning and intervention optimisation.

Computational cardiac biomechanics helps bridge this gap by combining physics-based models with anatomy derived from patient images [1–3]. Within this paradigm, image-driven FEA has become a widely used framework for modelling cardiac mechanics in individual subjects. Typically, ventricular geometry reconstructed from imaging data is incorporated into simulations to estimate full-field biomechanical fields under physiological loading conditions [4].

This approach has been applied to both healthy and pathological hearts, providing quantitative estimates of myocardial deformation, stress distribution, and indices of contractile function that complement conventional imaging assessments [4–6]. For example, FEA models have been used to analyse myofibre stress and strain patterns during LV remodelling and to quantify mechanical differences between infarcted and remote myocardium [4, 7]. Such analyses provide a mechanistic perspective on cardiac disease processes beyond what can be inferred from imaging alone.

Beyond fundamental studies, image-based FEA has also been explored in clinically oriented applications. By incorporating subject-specific geometry together with image-derived motion and boundary conditions, simulations can estimate regional myocardial mechanics that are difficult to measure directly using routine clinical imaging [4]. These mechanical indicators have been investigated in conditions such as ischaemic LV remodelling following myocardial infarction, hypertrophic cardiomyopathy associated with altered fibre mechanics, and LV aneurysm characterised by regional wall thinning [4, 7]. In interventional settings, including valve repair or replacement and other structural heart procedures, image-based computational models can be used to simulate tissue deformation and device–tissue interaction under physiological loading, thereby supporting pre-procedural planning, prediction of intervention outcomes, patient selection, and optimisation of intervention strategies and device configuration [4, 8].

1.1.2 Computational Challenges in Personalised Cardiac FEA

Despite its broad utility and demonstrated clinical relevance, personalised cardiac FEA remains constrained by substantial computational demands [9, 10]. Accurate simulation of ventricular mechanics typically requires nonlinear myocardial constitutive behaviour, large-deformation kinematics, and finely resolved meshes to represent complex anatomy and heterogeneous tissue properties [11]. These requirements lead to large-scale nonlinear systems and iterative solution procedures,

such that a single forward simulation may require many hours of computation, even when executed on modern High-Performance Computing (HPC) resources [10].

The computational burden increases further in inverse or optimisation-based workflows. Estimation of the unloaded or zero-pressure reference configuration, for example, commonly relies on iterative procedures that invoke repeated forward solves and enforce stringent convergence under near-incompressible material behaviour [12, 13]. Parameter identification, sensitivity analysis, and uncertainty quantification similarly require multiple evaluations of the forward model, compounding overall runtime [13]. Hence, personalised cardiac FEA is difficult to perform within clinically acceptable timescales, particularly in scenarios that require rapid feedback or interactive analysis [14].

Beyond solver cost, the end-to-end modelling pipeline introduces additional practical challenges. Traditional image-based cardiac FEA involves multiple sequential steps, including medical image segmentation to extract relevant anatomical structures, image registration to establish consistent spatial correspondence across cardiac phases or imaging modalities, geometry reconstruction, mesh generation and quality control, specification of material parameters and boundary conditions, execution of multi-cycle simulations to ensure periodic convergence, and post-processing of high-dimensional outputs [9, 15]. Several stages of this pipeline, especially registration, segmentation, and geometry preparation, remain semi-manual and time-consuming, particularly when image quality is limited or when dense meshes are required for numerical accuracy [9]. Therefore, many studies are restricted to relatively small cohorts, limiting statistical power and hindering large-scale or high-throughput investigations [15].

In addition to computational expense, the reliance on specialised expertise in continuum mechanics, numerical methods, and FEA software constitutes a further barrier to adoption. Model construction, parameter tuning, and interpretation of simulation outputs typically require substantial technical training, which restricts accessibility for non-expert users and complicates integration into routine clinical or translational workflows [15].

These limitations motivate the need for efficient alternatives to traditional FEA workflows, whereby the entire modelling pipeline can be greatly accelerated while maintaining sufficient accuracy, thus enabling broader application in both research and clinical practice [9].

1.1.3 DL for Image-driven Biomechanics

DL techniques have been increasingly applied to image-driven cardiac biomechanics because they can learn complex relationships directly from data, often with substantially faster inference and reduced reliance on manual optimisation compared to traditional approaches. Advances in computational hardware and large annotated datasets have enabled DL models to extract anatomical and motion information from medical images effectively, thereby alleviating some of the computational bottlenecks in personalised modelling workflows.

In cardiac image analysis, DL-based methods have become widely adopted for segmentation and registration of ventricular and myocardial structures from modalities including MRI, ECHO, and Computed Tomography (CT), often achieving high accuracy with substantially reduced processing time compared with traditional model-based techniques [16]. However, both segmentation and registration models still face challenges in generalising across different imaging modalities and acquisition conditions, and relatively few approaches have been demonstrated for robust 4D (3D+t) registration of cardiac sequences in realistic clinical settings [17].

In addition to image analysis, DL has been explored to accelerate FEA through surrogate modelling, where neural networks are trained to approximate the mapping from anatomical inputs and boundary conditions to biomechanical outputs. A range of data-driven surrogate architectures, including graph-based designs, have been proposed and have demonstrated substantial reductions in computation time for passive mechanical response prediction [18–20]. However, surrogate modelling of active contraction remains limited, partly due to the additional modelling complexity and the scarcity of large-scale, high-quality FEA datasets that incorporate active stress and full-cycle dynamics. On the other hand, physics-informed approaches that embed governing equations into the learning process are also a popular strategy for constructing FEA surrogates. Although some studies have attempted to, and in certain cases successfully, incorporate active contraction, these methods still face challenges in handling realistic complex geometries and capturing high-frequency spatio-temporal behaviour at scale [21, 22].

Overall, there is currently no universally robust and widely adopted solution, in either academia or industry, that provides a complete end-to-end replacement for each stage of the image-driven biomechanics pipeline. The limitations of existing DL approaches in terms of dataset dependency, geometric and domain generalisability, and comprehensive mechanical representation motivate the focus of this thesis on two main aspects of image-driven cardiac biomechanics. The first focus is on advancing DL-based image registration, with particular emphasis on robust 4D registration of

dynamic cardiac image sequences. The second focus is on developing comprehensive DL surrogate frameworks to accelerate FEA while capturing full-cycle dynamics, thereby enabling more efficient and scalable personalised biomechanics modelling for both research and clinical applications.

1.2 Research Aims

The ultimate goal of this research is to **realise a generalised DL-based surrogate framework that can effectively replace key components of traditional FEA in image-driven cardiac biomechanics, achieving substantial improvements in computational efficiency while preserving biomechanical fidelity**. To achieve this overarching objective, this thesis develops, improves, and optimises several key stages of the personalised modelling pipeline using DL techniques. The research is structured around four specific aims, focusing on the LV myocardium:

1. Cross-Modal Cardiac Motion Tracking

To develop robust DL-based methods for cross-modal image registration in both 2D+t and 3D+t cardiac sequences, enabling consistent spatial alignment of dynamic cardiac anatomy across different imaging modalities and time frames.

2. Physics-Informed Reduced-Order Mechanics Inference

To design and implement a reduced-order surrogate modelling framework that can autonomously infer patient-specific material parameters from selected time frames and subsequently provide rapid predictions of full-cycle ventricular mechanical responses.

3. Fast Prediction of the Unloaded Reference State

To develop a DL model for rapid prediction of the zero-pressure (unloaded) reference configuration mesh, which is a critical initial state for accurate FEA simulations but is traditionally obtained through iterative and computationally expensive inverse procedures.

4. Pre-Trained Full-Cycle Cardiac Mechanics Surrogate

To construct a pre-trained, comprehensive surrogate framework that does not rely on reduced-order modal representations and is capable of near-real-time prediction of full-cycle cardiac mechanics, including active stress behaviour, across diverse patient geometries and physiological conditions.

1.3 Thesis Outline

This thesis is organised into seven chapters:

Chapter 2 reviews prior work on cardiac image registration and image-driven FEA for LV myocardium, and summarises the methodological foundations that underpin the remainder of the thesis.

Chapter 3 presents a DL framework for cross-modal cardiac image registration in both 2D+t and 3D+t settings, incorporating a B-spline formulation with Fourier-domain regularisation to improve spatio-temporal consistency (Aim 1).

Chapter 4 develops a reduced-order physics-informed surrogate for LV mechanics that infers subject-specific material parameters from selected temporal frames and enables efficient prediction of full-cycle mechanical responses (Aim 2).

Chapter 5 proposes a graph-based model to rapidly estimate the zero-pressure (unloaded) reference configuration, providing an efficient alternative to conventional iterative procedures (Aim 3).

Chapter 6 introduces a pre-trained graph-based surrogate framework that does not rely on reduced-order modal representations and enables near-real-time prediction of full-cycle cardiac mechanics, including active stress behaviour (Aim 4).

Chapter 7 concludes with a summary of the main contributions, a discussion of limitations and implications, and an outlook on future research directions.

1.4 Publications and Conferences

This section summarises the journal articles and conference contributions produced during my doctoral research.

Journal papers

- S. Mu et al., “Imc-pinn-fe: A physics-informed neural network for patient-specific left ventricular finite element modeling with image motion consistency and biomechanical parameter estimation,” *arXiv preprint arXiv:2506.20696*, 2025 **(under review)**

- S. Mu et al., “Heartunloadnet: A cycle-consistent graph network with reduced supervision for predicting unloaded cardiac geometry from diastolic states,” *Computer Methods and Programs in Biomedicine*, p. 109 243, 2026
- S. Mu et al., “A cycle-consistent graph surrogate for full-cycle left ventricular myocardial biomechanics,” *arXiv preprint arXiv:2602.06884*, 2026 **(under review)**

Conferences (oral presentation)

- S. Mu et al., “An unsupervised deep learning framework for full cycle cardiac motion tracking in cross-modality images,” in *29th Congress of the European Society of Biomechanics (ESBiomech 2024): Abstracts*, 2024
- S. Mu et al., “A PINN network for patient-specific left ventricular finite element modeling with image motion-consistency,” in *30th Congress of the European Society of Biomechanics (ESBiomech 2025): Abstracts*, 2025
- S. Mu et al., “A physics-informed neural network for patient-specific left ventricular finite element modeling with image motion-consistency and myocardial stiffness and active tension estimation,” in *ASME SB3C Summer Bioengineering Conference (SBC 2025) Proceedings Book: Abstracts*, 2025

Related Work

Contents

2.1 Introduction	8
2.2 Cardiac Image Registration and Motion Tracking	9
2.2.1 2D Tracking Techniques	9
2.2.2 3D Tracking Techniques	12
2.3 Finite Element Analysis of Left Ventricular Mechanics	14
2.3.1 Principles	14
2.3.2 Clinical use of FEA	17
2.4 ML Surrogates in Cardiac Biomechanics	19
2.4.1 Early ML Surrogates	19
2.4.2 PINN-based Surrogates	20
2.4.3 Graph-based Surrogates	25

2.1 Introduction

This chapter reviews the background and methodological foundations underlying the subsequent chapters of this thesis. It first introduces cardiac image registration and motion tracking techniques for both 2D+t and 3D+t data, followed by an overview of the fundamental principles of FEA used throughout this work, together with representative examples of its clinical application. Finally, it reviews recent advances in machine learning surrogate models for cardiac biomechanics, highlighting both their potential to accelerate biomechanical simulation and the remaining challenges in generalisation, physical consistency, and clinical applicability.

2.2 Cardiac Image Registration and Motion Tracking

Cardiac image registration and motion tracking aim to quantify ventricular or myocardial deformation from time-resolved medical images and are fundamental to image-based cardiac analysis. In recent years, DL-based methods for 2D cardiac image registration have reached a mature stage in pairwise settings, with many approaches achieving accurate and robust pairwise alignment. However, most methods remain frame-wise, with only limited work enforcing temporal or spatio-temporal consistency to model cardiac motion over an entire cycle. This is a non-trivial limitation, as cardiac motion is inherently continuous and cyclic, and temporally consistent deformation fields are essential for both image-based analysis and downstream biomechanical FEA modelling. By contrast, 3D and 3D+t registration is less developed due to data and modelling constraints, and spatio-temporally consistent approaches are still emerging. Accordingly, this section reviews representative registration methods across 2D, 2D+t, 3D, and 3D+t settings, covering both traditional and learning-based techniques, to provide background for the motion-consistent frameworks presented in Chapter 3.

2.2.1 2D Tracking Techniques

Non-DL methods: Before the surge of DL, 2D cardiac tracking was mainly developed within classical image processing and optimisation frameworks. In practice, most pipelines fell into three closely related forms: pattern-based tracking, feature or contour tracking, and intensity-driven non-rigid registration.

Pattern-based tracking, typified by speckle tracking and block matching in ECHO, estimates displacement by assuming local pattern constancy across frames. These methods are attractive because they are simple and computationally efficient, but they are sensitive to decorrelation, large deformation, and out-of-plane motion when anatomical structures leave the imaging plane [29, 30]. Closely related approaches include optical flow methods, which estimate dense motion from the brightness constancy assumption. Depending on the formulation, optical flow may impose global smoothness, as in the Horn and Schunck model, or estimate local motion within a neighbourhood, as in the Lucas and Kanade formulation [31, 32]. Optical flow has therefore been widely investigated for myocardial motion analysis in cardiac ultrasound [33]. However, in practice, the brightness constancy assumption can be violated in real echocardiographic sequences because image appearance changes over time, reducing robustness unless additional regularisation or multiscale

estimation strategies are introduced [34].

2

Another family of approaches relies on feature or contour-based tracking, where motion is estimated by first extracting geometric entities and then matching them across time. Typical examples include shape-based boundary correspondence methods that track the LV endocardial contour point-wise over a cardiac cycle [35], as well as feature-based point tracking schemes that analyse the reliability of different features and the resulting motion feature decorrelation. When contours or landmarks are reliably available, these methods can produce accurate local motion estimates. However, their performance depends strongly on segmentation and feature detection quality, which can become unreliable in low-contrast or noisy imaging conditions [36].

A more general formulation casts motion tracking as a variational image registration problem. In this framework, the deformation field is obtained by minimising an energy functional that combines an image similarity term with spatial (and sometimes physiological) regularisation [36]. In practice, widely used optimisation-based solvers include B-spline free-form deformation models [37] and demons-style non-parametric registration methods together with their diffeomorphic variants [38]. These approaches can produce dense motion fields and are flexible across imaging modalities. However, they remain sensitive to initialisation and hyper-parameter selection and may converge to local minima, particularly under low-contrast or noisy imaging conditions. In addition, they typically require sequence-specific optimisation procedures [36].

Beyond purely pairwise 2D registration that aligns frames independently, a line of classical work introduced explicit temporal constraints and sequence-level (2D+t) formulations, where motion is estimated by jointly coupling all frames in time to reduce drift and enforce coherent trajectories [39, 40]. For example, Ledesma-Carbayo et al. proposed a spatio-temporal non-rigid registration approach that directly solves for a deformation field over space and time by minimising the difference to a reference frame, thereby reducing accumulated error compared with purely pairwise alignment [39]. De Craene et al. later introduced Temporal Diffeomorphic Free-Form Deformation (TDFFD), enforcing temporal consistency by representing a 4D velocity field as a sum of continuous spatio-temporal B-spline kernels and integrating this non-stationary field over time to recover a coherent displacement trajectory [40]. More recently, Wiputra et al. proposed a post-registration regularisation layer that fits a periodic motion model based on B-splines of Fourier to displacement fields obtained from pairwise registration, improving cycle-level smoothness without modifying the underlying pairwise optimiser [41, 42]. Overall, these spatio-temporal formulations tend to produce smoother and more physiologically plausible motion fields. However, they in-

introduce additional modelling and optimisation complexity and typically remain sequence-specific, which limits scalability in large-cohort studies and time-critical workflows [39, 40].

DL Methods: DL methods have become increasingly popular for 2D(+t) cardiac tracking because they replace sequence-specific optimisation with amortised inference, often providing substantially faster runtime while maintaining competitive accuracy in dense motion estimation [43]. A common starting point is supervised learning, where a Convolutional Neural Network (CNN) is trained to regress motion fields or related quantities from labelled data. In practice, however, dense motion labels are expensive, and segmentation-derived supervision can be unreliable in ECHO, which limits the scalability and robustness of purely supervised pipelines.

To reduce the dependence on labels, many approaches adopt unsupervised or semi-supervised learning based on differentiable warping. In this setting, the network predicts a displacement field that warps a moving frame to match a fixed frame, and training is driven by image similarity and regularisation rather than ground-truth flow, typically implemented via a differentiable spatial transformer layer [43, 44]. This formulation has been successfully adapted to cardiac ECHO. For example, Evain et al. proposed a DL motion estimator for 2D ECHO based on an optical flow style design and validated it with a realistic simulation pipeline to improve robustness under clinical image appearance variations [45]. Semi-supervised variants further exploit limited annotations by coupling tracking with auxiliary tasks. Ta et al. introduced a joint learning framework that links motion estimation and segmentation during training, using the estimated motion to propagate labels across time and thereby strengthening supervision under sparse manual annotations [46].

Beyond basic pairwise estimation, recent work has started to address two practical deployment issues, domain shift and temporal drift, and both tend to be amplified when tracking or registration is performed across modalities. For domain shift, FOAL introduces a fast online adaptation mechanism that updates a pre-trained tracker on a new test sequence using only a small amount of test-time optimisation, improving robustness when the test distribution differs from the training data [47]. For temporal drift, sequence-consistent designs explicitly link motion across time. DeepTag is an example for tagged MRI, where inter-frame motion is estimated and then composed through a differentiable layer to recover reference-to-any-frame trajectories over the full sequence [48]. When moving to cross-modality, the challenge further increases because the same anatomy can exhibit substantial appearance and contrast differences, which weakens photometric losses and intensity-constancy assumptions. This has motivated methods that explicitly learn inter-modality correspondences, such as cross-modal attention for multi-modal registration [49].

However, extending temporally consistent designs to cross-modality 2D+t tracking remains challenging in practice, because large appearance gaps reduce the reliability of photometric supervision and paired time-resolved multi-modal data are less commonly available [49].

2.2.2 3D Tracking Techniques

Compared with 2D tracking, 3D and 3D+t motion tracking estimates deformation in volumetric space, which is essential for recovering through-plane motion and obtaining physically meaningful 3D strain patterns. In practice, 3D tracking remains less mature than 2D due to (i) higher acquisition cost and smaller publicly available datasets, (ii) lower temporal resolution and stronger noise/artifacts in volumetric imaging (especially 3D echocardiography), and (iii) the increased ambiguity of the motion field when image information is sparse or anisotropic across slices.

Non-DL methods. Early 3D tracking pipelines were largely dominated by optimisation-based volumetric registration, where a dense 3D deformation field is estimated between two frames by minimising an image-matching objective with spatial regularisation. In 3D echocardiography, a representative class of methods improves physiological plausibility by adapting the transformation model to cardiac anatomy. For example, anatomical free-form deformation (AFFD) organises B-spline control points (and the associated smoothness regularisation) in heart-aware coordinates, so that smoothing is applied along more physiological directions and 3D strain estimation becomes more robust [50, 51]. Multi-view strategies address view-dependent artefacts by estimating a consistent transformation directly from multiple 3D ultrasound views rather than from a single compounded volume, which improves robustness when individual views suffer from shadowing, dropouts, or limited field-of-view [52]. For cardiac MRI, biomechanical deformable models embed motion in a volumetric mechanics-based representation and use this as a strong prior to regularise deformation over time, supporting consistent 4D (3D+t) analysis [53].

Despite these advances, 3D pairwise registration remains computationally expensive, sensitive to initialisation and hyper-parameters, and its frame-wise nature can introduce drift and frame-to-frame jitter when applied across a full cardiac cycle. To recover a temporally coherent 3D+t trajectory without redesigning the underlying 3D optimiser, a practical strategy is post-registration temporal regularisation. After pairwise registration, a periodic spatio-temporal motion model based on B-splines or Fourier can be fitted to the displacement fields to enforce cycle-level smoothness and periodicity and suppress drift [41, 42]. Overall, non-DL pipelines are interpretable and

can produce physiologically meaningful motion, but robust 3D+t tracking typically requires either explicit temporal modelling (e.g., biomechanical 4D formulations) or an additional temporal coherence stage (e.g., B-spline Fourier fitting).

DL methods. Recent DL approaches for 3D motion tracking are mainly unsupervised or weakly supervised, because voxel-wise ground-truth 3D motion is rarely available in clinical data. Most methods therefore learn deformation fields through differentiable warping, combining image reconstruction/similarity losses with spatial regularisation, and, when possible, temporal-consistency constraints to stabilise trajectories across a cardiac cycle.

For cine Cardiovascular Magnetic Resonance (CMR), a major practical issue is that routine acquisitions are often multi-slice 2D (short-axis and long-axis) rather than true 3D volumes, which makes through-plane motion poorly constrained. Multi-view learning addresses this by fusing orthogonal views to infer a consistent 3D motion field. MulViMotion is a representative example: it integrates short-axis and long-axis cine CMR within a hybrid 2D/3D network and introduces shape-related weak supervision to improve 3D consistency, thereby reducing through-plane ambiguity [54]. The remaining limitations are that performance depends on the availability and quality of multi-view acquisitions, and generalisation can degrade under scanner/protocol/pathology shifts.

For volumetric ultrasound, learning-based 3D+t tracking is further challenged by speckle noise, shadowing and dropouts, and the lack of reliable annotations. Recent work has therefore explored end-to-end 4D (3D+t) motion estimation with temporal constraints. For example, unsupervised spatiotemporal optical-flow style networks combine a spatial reconstruction objective with an explicit temporal-consistency loss to exploit pairwise and temporal correlations and reduce drift in 4D echocardiography sequences [55]. Attention-based designs have also been studied for 3D echocardiography motion/strain analysis to improve inter-frame feature matching and robustness under noise [56]. In practice, a common weakness of current 3D+t DL pipelines is that validation is often limited by dataset scale and ground-truth availability (frequently relying on synthetic or weak supervision), and temporal stability can still degrade in regions with severe artefacts or low texture.

Overall, 3D(+t) tracking remains constrained by data availability and modality-specific limitations, but the trend is toward spatio-temporally consistent learning frameworks that either exploit multi-view cine CMR priors or incorporate temporal regularisation directly for 3D ultrasound motion estimation. These developments provide the methodological basis for the 3D motion-consistent designs considered in this thesis.

2.3 Finite Element Analysis of Left Ventricular Mechanics

2.3.1 Principles

FEA has become a well-established framework for modelling LV biomechanics, with numerous implementations reported in the literature [57–61]. In such models, myocardial mechanics are typically described using a hyperelastic constitutive law, most commonly the Holzapfel–Ogden model [62] or a transversely isotropic Fung-type formulation [63]. Active contraction is incorporated through an additional active stress component, often represented using phenomenological models such as the Guccione model [63] or biophysically motivated formulations such as the Niederer model based on sarcomere-level calcium dynamics [64]. In this thesis, the Guccione model was adopted due to its computational efficiency and widespread use in cardiac biomechanics literature.

To represent myocardial fibre architecture, the model usually assumes a transmurally varying helix angle, which can be obtained by solving a Laplace equation with Dirichlet boundary conditions prescribed on the endocardial and epicardial surfaces. For circulatory coupling, the ventricular model may also be linked to a lumped-parameter representation of the cardiovascular system to account for pressure and flow interactions between the ventricles and the great arteries.

For spatial discretisation, the myocardium is typically meshed using either first-order or second-order finite elements. First-order elements generally require approximately 10^5 – 10^7 elements to achieve mesh convergence [65, 66], whereas second-order elements can reach similar accuracy with approximately 10^3 – 10^4 elements [67, 68].

In this section, we provide details of one representative implementation described in [57], which forms the LV mechanics setup used throughout the following chapters.

The passive strain energy density W_{passive} is modelled using a fully incompressible, transversely isotropic Fung-type hyperelastic law with a transmural helix-angle fibre architecture [69, 70]:

$$W_{\text{passive}} = \frac{1}{2} C (\exp(Q) - 1), \quad (2.1)$$

$$Q = b_f E_{ff}^2 + b_{xx} (E_{ss}^2 + E_{nn}^2 + 2E_{sn}^2) + 2b_{fx} (E_{fs}^2 + E_{fn}^2), \quad (2.2)$$

where $\mathbf{E}$ is the Green–Lagrange strain tensor,

$$\mathbf{E} = \frac{1}{2} (\mathbf{C} - \mathbf{I}), \quad \mathbf{C} = \mathbf{F}^T \mathbf{F}, \quad \mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}}, \quad (2.3)$$

where $E_{ij} = \mathbf{i}_0^T \mathbf{E} \mathbf{j}_0$ are the strain components expressed in the local orthonormal basis $\{\mathbf{f}_0, \mathbf{s}_0, \mathbf{n}_0\}$, denoting the fibre (**f**), sheet (**s**), and sheet-normal (**n**) directions. The coefficients b_f , b_{xx} , and b_{fx} define the anisotropic stiffness, and C is a global stiffness scaling factor. Here $\mathbf{F}$ is the deformation gradient tensor, while $\mathbf{x}$ and $\mathbf{X}$ denote the deformed and reference coordinates, respectively.

We assume a linear transmural helix-angle variation

$$\theta_h = \theta_{\text{endo}} + \xi(\theta_{\text{epi}} - \theta_{\text{endo}}),$$

where $\xi \in [0, 1]$ is a transmural coordinate ($\xi = 0$ endocardium, $\xi = 1$ epicardium). A scalar field $\phi(\mathbf{X})$ is defined on the reference mesh with $\phi = 0$ on the endocardium and $\phi = 1$ on the epicardium, and we set $\xi = \phi$. The sheet-normal direction is obtained from the reference gradient,

$$\mathbf{n}_0 = \frac{\nabla_0 \phi}{\|\nabla_0 \phi\|}.$$

Let $\mathbf{e}_c$ and $\mathbf{e}_l$ denote local circumferential and longitudinal unit directions (tangent to the ventricular wall). An in-plane fibre direction is defined as

$$\tilde{\mathbf{f}}_0 = \cos \theta_h \mathbf{e}_c + \sin \theta_h \mathbf{e}_l,$$

which is orthogonalised against $\mathbf{n}_0$ and normalised to obtain $\mathbf{f}_0$. The sheet direction is then defined as

$$\mathbf{s}_0 = \frac{\mathbf{n}_0 \times \mathbf{f}_0}{\|\mathbf{n}_0 \times \mathbf{f}_0\|}.$$

The active fibre tension T_{active} is determined by a calcium activation model [69]:

$$T_{\text{active}} = T_{\text{max}} \frac{Ca_0^2}{Ca_0^2 + ECa_{50}^2} C_t, \quad (2.4)$$

where T_{max} is the maximum active tension, Ca_0 is the peak calcium concentration, ECa_{50} is the sarcomere length-dependent calcium sensitivity, and C_t captures temporal activation:

$$ECa_{50} = \frac{(Ca_0)_{\text{max}}}{\sqrt{\exp(B(l - l_0)) - 1}}, \quad C_t = \frac{1}{2} (1 - \cos \omega). \quad (2.5)$$

Here B is a constant, $(Ca_0)_{\text{max}}$ is the maximum peak intracellular calcium concentration, and l_0 is the sarcomere length at which no active tension develops. The phase ω depends on the cardiac

cycle time:

$$\omega = \begin{cases} \pi \frac{t}{t_0} & 0 \leq t < t_0, \\ \pi \frac{t - t_0 + t_r}{t_r} & t_0 \leq t < t_0 + t_r, \\ 0 & t_0 + t_r \leq t, \end{cases} \quad (2.6)$$

where t_0 is the time to peak tension and t_r is the relaxation time [69],

$$t_r = ml + b, \quad (2.7)$$

with m the slope and b the intercept of the linear relaxation-duration-sarcomere-length relation.

The sarcomere length l varies with fibre stretch λ :

$$\lambda = \sqrt{\mathbf{f}_0 \cdot \mathbf{C} \mathbf{f}_0}, \quad l = \lambda l_r, \quad (2.8)$$

where l_r is the relaxed sarcomere length.

The passive Cauchy stress is obtained from the strain energy density via [71]

$$\boldsymbol{\sigma}_{\text{passive}} = \frac{2}{J} \mathbf{F} \frac{\partial W_{\text{passive}}}{\partial \mathbf{C}} \mathbf{F}^T, \quad J = \det(\mathbf{F}), \quad (2.9)$$

and the active Cauchy stress is applied in the current fibre direction

$$\mathbf{f} = \frac{\mathbf{F} \mathbf{f}_0}{\|\mathbf{F} \mathbf{f}_0\|}, \quad \boldsymbol{\sigma}_{\text{active}} = T_{\text{active}} \mathbf{f} \otimes \mathbf{f}. \quad (2.10)$$

The total Cauchy stress is

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_{\text{passive}} + \boldsymbol{\sigma}_{\text{active}} - p \mathbf{I}, \quad (2.11)$$

where p is the Lagrange multiplier (hydrostatic pressure) enforcing incompressibility.

The FEA solves for the displacement field $\mathbf{u} = \mathbf{x} - \mathbf{X}$ and the cavity pressure P_{cavity} by enforcing quasi-static equilibrium in the deformed myocardial domain Ω :

$$\nabla \cdot \boldsymbol{\sigma} = \mathbf{0} \quad \text{in } \Omega, \quad (2.12)$$

subject to the traction boundary condition

$$\boldsymbol{\sigma} \mathbf{n} = -P_{\text{cavity}} \mathbf{n} \quad \text{on } \Gamma_{\text{endo}}, \quad (2.13)$$

and the incompressibility constraint $J = 1$, where $\mathbf{n}$ denotes the outward unit normal and Γ_{endo} is the endocardial surface.

In practice, this system is solved in weak form through the virtual work statement coupled with the incompressibility constraint:

$$\int_{\Omega_0} \mathbf{P} : \nabla_0 \delta \mathbf{u} \, dV_0 - \int_{\Gamma_{0,\text{endo}}} P_{\text{cavity}} J \mathbf{F}^{-T} \mathbf{N} \cdot \delta \mathbf{u} \, dA_0 = 0, \quad (2.14)$$

$$\int_{\Omega_0} \delta p (J - 1) \, dV_0 = 0, \quad (2.15)$$

where $\delta \mathbf{u}$ denotes the virtual displacement field, δp is the virtual pressure field associated with the incompressibility constraint, Ω_0 is the reference myocardial domain with volume element dV_0 , $\Gamma_{0,\text{endo}}$ is the endocardial surface with surface element dA_0 , $\mathbf{N}$ is the outward unit normal in the reference configuration, ∇_0 denotes the gradient operator with respect to the reference coordinates $\mathbf{X}$, and $\mathbf{F} = \mathbf{I} + \nabla_0 \mathbf{u}$. The basal plane was constrained to suppress rigid-body motion while allowing physiologically realistic ventricular deformation. The first Piola–Kirchhoff stress $\mathbf{P}$ is related to the Cauchy stress by

$$\mathbf{P} = J \boldsymbol{\sigma} \mathbf{F}^{-T}.$$

2.3.2 Clinical use of FEA

In clinical applications of cardiac mechanics, FEA typically yields quantitative outputs such as P-V loops, ventricular pressure traces, and myocardial stress trajectories, which can be used to predict intervention response and compare candidate treatment strategies. For example, as shown in Fig. 2.1, Green et al used patient-specific biomechanical outputs, including P-V loops and myofibre stress trajectories, to distinguish fetal cases that later achieved biventricular versus univentricular outcomes after fetal aortic valvuloplasty. This study showed that pre-intervention peak systolic myofibre stress provided stronger discrimination than conventional echocardiographic indices, highlighting the value of FEA-derived biomechanical markers for patient stratification [57].

Moreover, as shown in Fig. 2.2, Fan et al used a cardiac FEA electromechanics–perfusion

model to compare candidate pacing sites for CRT. CRT is a treatment for heart failure in which electrical pacing is used to re-coordinate ventricular contraction, and its benefit depends in part on where the pacing lead is placed. In this figure, the model compares different candidate sites using changes in left ventricular pressure, activation time, the rate of pressure rise (dP/dt), and ejection fraction. A larger improvement in pressure generation and ejection fraction, together with a greater reduction in activation time, indicates a more favourable pacing location. These outputs show that the haemodynamic and mechanical response varies across candidate sites, allowing the model to identify more favourable pacing locations for CRT planning [72].

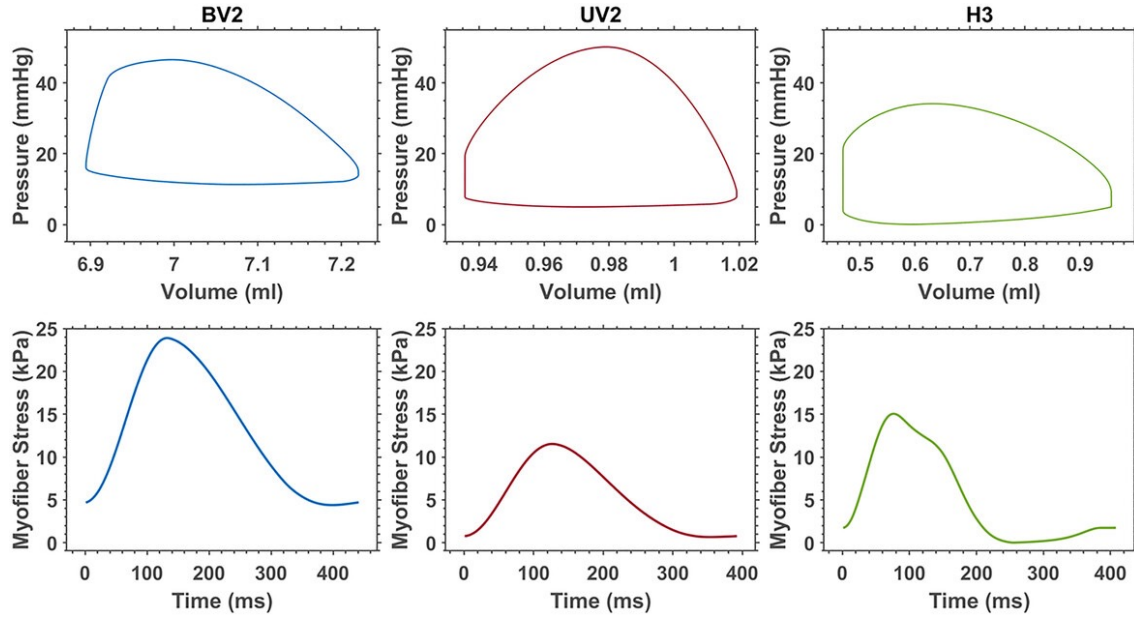


Figure 2.1: **Representative biomechanical outputs from patient-specific FEA in fetal cardiac intervention.** The top row shows LV P-V loops, and the bottom row shows the corresponding myofiber stress trajectories over the cardiac cycle, for three representative fetal cases: BV2, UV2, and H3. BV2 and UV2 denote cases with later biventricular and univentricular outcomes, respectively, while H3 denotes a healthy control. The BV and UV cases exhibit distinct P-V behaviour and markedly different peak systolic myofiber stress, with the BV case showing a substantially higher stress peak than the UV case. These differences are consistent with the study finding that pre-intervention peak systolic myofiber stress is a strong predictor of fetal aortic valvuloplasty outcome [57].

These examples demonstrate that image-driven cardiac FEA can provide clinically relevant biomechanical and haemodynamic indicators for intervention planning and treatment optimisation. Beyond ventricular biomechanics alone, cardiac modelling has increasingly expanded toward multi-physics whole-heart digital twin frameworks integrating electrophysiology, haemodynamics, perfusion, and fluid-structure interaction. Large-scale initiatives such as the Living Heart Project have demonstrated the feasibility of combining anatomical, electromechanical, and circulatory modelling within unified computational frameworks for therapy planning, device evaluation, and

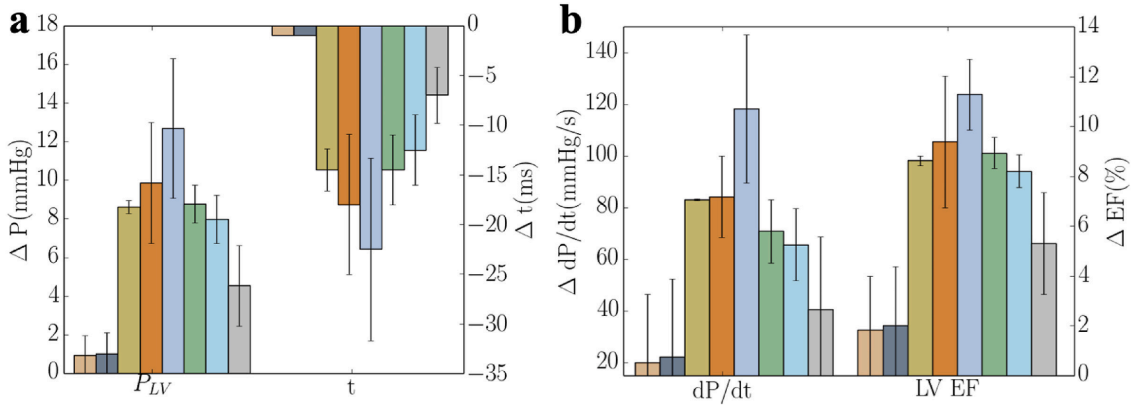


Figure 2.2: **FEA comparison of candidate pacing sites for CRT.** Panel (a) shows the predicted changes in left ventricular pressure (ΔP_{LV}) and activation time (Δt), while panel (b) shows the predicted changes in the maximum rate of pressure rise ($\Delta dP/dt$) and left ventricular ejection fraction (ΔEF). Each coloured bar represents one candidate pacing location evaluated in the model [72].

patient-specific simulation [73, 74]. Such frameworks have also been extended to incorporate valve dynamics and fluid–structure interaction, further improving the physiological realism of computational cardiac simulations. Nevertheless, despite these advances, the high computational cost and workflow complexity of multi-physics cardiac simulations continue to limit their scalability and broader clinical translation.

2.4 ML Surrogates in Cardiac Biomechanics

Significant advancements have been made in the use of ML models as surrogates for FEA simulations in cardiovascular mechanics. These methodologies can be broadly divided into three phases: early ML surrogates, PINN-based strategies, and emerging graph-based surrogates.

2.4.1 Early ML Surrogates

In the early development phase of FEA surrogate modelling for cardiovascular biomechanics, researchers turned to standard machine-learning methods, such as MLPs or gradient-boosted trees. For example, Liang L. et al. used an MLP to predict aortic wall-stress distributions, thereby providing a fast and accurate FEA surrogate [75]. The innovation of this work was in demonstrating that a data-driven model could reproduce spatial stress fields rather than simply scalar outputs, yet it was limited by the fact that the geometry, material properties, and loading conditions were restricted to a narrow domain and the model required a fixed mesh topology and

consistent geometry preprocessing (e.g., PCA encoding of node positions). As research shifted toward more complex cardiac structures, especially the LV, Dabiri Y. et al. applied XGBoost to capture both passive and active LV mechanics, achieving substantial increases in simulation speed and flexibility over traditional FEA models [76]. Although this represented an advance toward coupling active and passive mechanics and predicting P-V loops, the model was still confined to the parameter space sampled during training and lacked true generalisation to new geometries or altered loading conditions. In 2020, a further step was made, first, an MLP-based surrogate was developed that simplified the input domain and rapidly simulated LV deformation under variable geometries and material-loading combinations during the diastolic phase [77], and later that year, an approach combining feed-forward and recurrent neural networks (LSTM + MLP) used clinical-style parameters to predict LV kinematics directly, entirely bypassing mesh input [78]. In parallel, several studies from 2018 onward explored Gaussian Process (GP)-based statistical emulators for uncertainty-aware parameter inference and acceleration of computationally expensive LV FEA simulations [79–81]. However, these methods remained primarily suitable for low-dimensional parameter spaces and limited geometric variability.

Moreover, ML surrogates have also been applied to inverse FEA modelling. For example, Cai et al. employed K-nearest neighbour, XGBoost and MLP-based surrogates to estimate LV myocardial parameters rapidly from P-V and strain data [82]. Similarly, Babaei et al. developed an ML model that uses ED P-V relationship together with geometry and fibre-orientation features to inversely estimate passive myocardial stiffness [83]. In a more recent study, Chen Y. et al. [84] developed a ResNet18-based deep-learning model that predicts myocardial material parameters directly from routine cardiac MRI data, achieving mean relative errors under 5% for key parameters.

While these methods delivered dramatic speed-ups and reduced reliance on detailed meshes, their limitations remained. Training datasets still spanned narrow parameter domains, model outputs remained largely single-task, and thus the frameworks were still not yet ready for the more complex, heterogeneous clinical cases with diverse anatomies and loading states.

2.4.2 PINN-based Surrogates

2.4.2.1 Overview of PINN in cardiac mechanics

Beyond early vanilla MLP-based approaches, unsupervised and weakly supervised methods have become increasingly popular because they can be trained without fully pre-labelled data and can

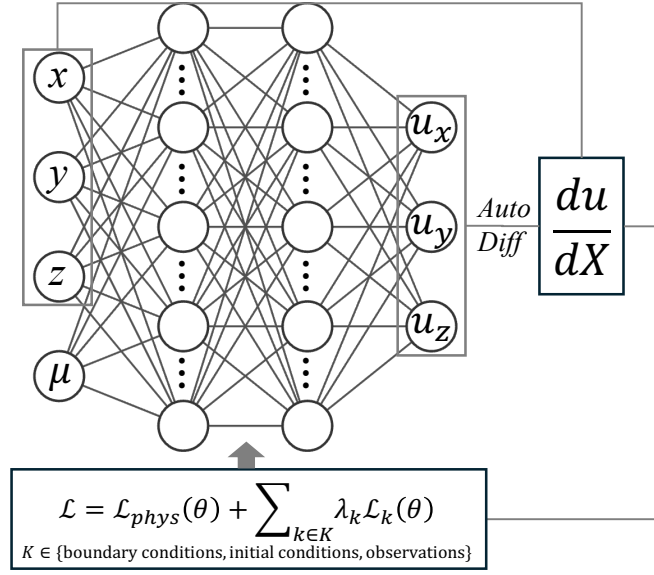


Figure 2.3: **Vanilla PINN schematic in cardiac mechanics.** Spatial coordinates and problem-related parameters are provided as inputs to a FCNN, which predicts displacement fields throughout the computational domain. AD is then used to compute spatial derivatives required by the governing equations. The network is trained by minimising a composite loss function consisting of physics-based residuals together with boundary conditions, initial conditions, and observational constraints.

more easily adapt to different material parameters. A representative family is the PINN. Several recent review papers have summarised the development of PINNs and physics-informed scientific machine learning across a broad range of engineering and biomedical applications [85–87]. Unlike purely data-driven models, a PINN enforces the governing Partial Differential Equation (PDE) of continuum mechanics, together with boundary and initial conditions, in a way that is conceptually similar to the constraints used in FEA. The key idea is that the spatial derivatives needed by the PDEs are computed using AD applied to the network outputs, allowing the governing laws and constraints to be written as differentiable loss terms [88].

In cardiac biomechanics, most PINNs are not pre-trained. Instead, they are commonly used in a patient-specific setting, where the network parameters are directly optimised for a single anatomy using the available measurements, making these approaches conceptually closer to optimisation-based parameter estimation than conventional pre-trained machine learning models. In this case, the network inputs can be kept simple and are not required to encode every factor explicitly, which helps reduce model complexity. Despite this per-case optimisation, PINNs can still be much faster than repeated inverse FEA procedures because they rely on efficient gradient-based optimisation and can exploit modern hardware acceleration. However, their performance depends strongly on problem setup (e.g., loss weighting, boundary conditions, and measurement quality), and they do

not yet consistently replace FEA across all clinical scenarios [85, 88, 89].

A vanilla PINN (Fig. 2.3) for cardiac biomechanics often estimates motion by parametrising the myocardial displacement field $\mathbf{u}$ with a neural network:

$$\mathbf{u} = \mathbf{u}((\mathbf{X}, \boldsymbol{\mu}); \boldsymbol{\theta}), \quad (2.16)$$

where $\mathbf{X} \in \Omega$ denotes the reference (unloaded) coordinate in the myocardial domain Ω , $\boldsymbol{\mu}$ represents additional physiological inputs such as material parameters, cavity volume, or pressure, and $\boldsymbol{\theta}$ are the trainable parameters.

Using AD, the PINN-predicted displacement field provides the spatial derivatives required to construct the deformation gradient:

$$\mathbf{F} = \mathbf{I} + \frac{\partial \mathbf{u}}{\partial \mathbf{X}}. \quad (2.17)$$

The network is trained by minimising a physics-based objective, optionally augmented with additional constraints depending on the available information (e.g., boundary conditions, near-incompressibility, or image-derived observations):

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \left(\mathcal{L}_{\text{phys}}(\boldsymbol{\theta}) + \sum_{k \in \mathcal{K}} \lambda_k \mathcal{L}_k(\boldsymbol{\theta}) \right), \quad \mathcal{K} \in \{\text{BC, IC, OBS}, \dots\}. \quad (2.18)$$

Here, $\boldsymbol{\theta}$ denotes the trainable neural-network parameters and λ_k are weighting coefficients balancing different loss components. $\mathcal{L}_{\text{phys}}$ enforces physical admissibility (e.g., equilibrium consistency), while the optional terms incorporate Boundary Condition (BC), Initial Condition (IC), and potential Observation (OBS) constraints, such as cavity-volume labels derived from images.

However, vanilla PINNs are often difficult to train for cardiac biomechanics because ventricular anatomy is geometrically complex and the mechanics are highly nonlinear. In particular, LV/Right Ventricle (RV) simulations involve large deformations, strong material anisotropy, near-incompressibility, and active contraction, all of which can make optimisation unstable and slow to converge. Despite these challenges, PINNs remain attractive because they can incorporate physical laws and reduce reliance on dense labels.

In the following, we summarise two representative PINN variants for LV modelling.

2.4.2.2 Representative Works in LV PINN

Although vanilla PINNs based on AD can work well for relatively simple PDE settings (e.g., WarpPINN for medical image tracking with an incompressibility constraint [90]), they often become difficult to train for FEA-like cardiac mechanics problems because the governing equations are highly nonlinear and stiff. To improve stability, most PINN–FEA approaches avoid computing spatial derivatives directly by AD on (x, y, z) . Instead, they adopt a FEA representation, where the displacement gradients are computed using FEA shape functions:

$$\nabla_0 \mathbf{u}_\theta = \mathbf{U}_\theta \nabla_0 \mathbf{H} \quad (2.19)$$

where $\mathbf{H}$ represents the FEA test/shape functions, $\mathbf{U}_\theta$ is the network-predicted nodal displacement vector, and $\nabla_0 \mathbf{u}_\theta$ is the displacement gradient with respect to the reference coordinates.

Under Eq. 2.19, once the nodal displacements are predicted, the displacement gradient is computed by applying the FEA shape-function derivatives. For a fixed reference tetrahedral mesh, these derivatives depend only on the reference geometry and can therefore be precomputed. This avoids using automatic differentiation with respect to spatial coordinates to obtain $\nabla_0 \mathbf{u}_\theta$, and the governing equations can be evaluated in the standard element-wise manner. In the following variants, linear 4-node tetrahedra are used, while efficient, they may be sensitive to mesh resolution and element quality near the apex [22]. The same framework can be extended to 10-node second-order tetrahedra in future work.

Motiwale et al.: Weak-Form Mechanics [21]

Motiwale et al. proposed Neural Network Finite Element (NNFE), a weak-form mechanics framework that predicts LV displacements from pressure over a full cardiac cycle. Compared with a vanilla PINN, NNFE introduces two key design differences.

First, it uses only endocardial pressure p and active tension T as inputs, while the global material stiffness, helix-angle field, and the unloaded reference coordinates $\mathbf{X}$ are predefined for each case and are not directly fed into the network. This substantially reduces input complexity because spatial coordinates (x, y, z) are not used as inputs, and it also offers flexibility across material parameters and FEA settings for patient-specific cases. In contrast, a pre-trained data-driven model may require very large datasets to cover diverse scenarios.

Second, it uses the weak form (integral form), as in FEA, rather than enforcing the strong-form

PDE pointwise. This replaces pointwise PDE residuals with an integrated force balance, which is typically more stable for complex ventricular mechanics. In NNFE, the displacement field is predicted as $\mathbf{u}_\theta(p, T)$, and the physics is enforced by driving the weak-form residual force vector toward zero over sampled input pairs $(p, T) \in \mathcal{D}$:

$$\mathcal{L}_{\text{phys}} = \sum_{(p, T) \in \mathcal{D}} \|\mathbf{R}(\mathbf{u}_\theta(p, T); p, T)\|_2^2. \quad (2.20)$$

Here $\mathbf{R}$ is the FEA-style “internal minus external” residual assembled from weak-form integrals:

$$\mathbf{R}(\mathbf{u}_\theta(p, T); p, T) = \mathbf{f}_{\text{int}}(\mathbf{u}_\theta(p, T); T) - \mathbf{f}_{\text{ext}}(p). \quad (2.21)$$

$$\mathbf{f}_{\text{int}} = \mathbf{f}_{\text{passive}} + \mathbf{f}_{\text{active}}. \quad (2.22)$$

$$\mathbf{f}_{\text{passive}} = \int_{\Omega_0} \mathbf{P}_{\text{pass}}(\mathbf{u}_\theta) : \nabla_0 \mathbf{H} \, d\Omega. \quad (2.23)$$

$$\mathbf{f}_{\text{active}} = \int_{\Omega_0} \mathbf{P}_{\text{act}}(\mathbf{u}_\theta, T) : \nabla_0 \mathbf{H} \, d\Omega. \quad (2.24)$$

$$\mathbf{f}_{\text{ext}}(p) = \int_{\partial\Omega_{N,0}} \mathbf{t}_p^0(p) \cdot \mathbf{H} \, d\Gamma. \quad (2.25)$$

where Ω_0 and $\partial\Omega_{N,0}$ denote the reference (unloaded) myocardial domain and its Neumann boundary, respectively. $\mathbf{H}$ represents the FEA test/shape functions used to assemble nodal forces, $\mathbf{P}$ is the first Piola–Kirchhoff stress, Γ denotes the endocardial surface, and $\mathbf{t}_p^0$ is the traction induced by the chamber pressure. Minimising $\mathcal{L}_{\text{phys}}$ therefore enforces equilibrium by matching the internal (passive + active) forces to the external pressure loading across the sampled (p, T) space.

After training, a P-V loop is generated by sampling a sequence of states $\{(p_k, T_k)\}_{k=1}^{N_t}$. For each sample, the trained NNFE model is evaluated to obtain the nodal displacement $\mathbf{u}_\theta(p_k, T_k)$. The mesh is then updated to the deformed configuration, and the cavity volume V_k is computed from the deformed endocardial surface, yielding the P-V loop $\{(V_k, p_k)\}_{k=1}^{N_t}$.

In the reported study, training (optimization) took approximately 2.25 hours, and a full P-V cycle was completed in about 30 seconds (versus ~ 5 hours for the FEA baseline), with a mean nodal displacement error of 0.023 ± 0.019 mm. A current limitation is that NNFE has mainly been demonstrated on an idealised LV geometry, extending it to realistic patient-specific ventricles with more complex anatomy and boundary conditions remains future work.

Buoso et al proposed another PINN variant to reduce complexity, the key is to transfer the output to reduced-order form. Specifically, the input is the same of Motiwale's NNFE (p and T), but output is the coefficients of motion modes for myocardium, which indeed reduces the dimensions and complexity of the overall network. The motion modes are generated using POD from 75 cases of Multi-Modal Whole Heart (MMWH) dataset [91–93]. With a settled reference geometry (unloaded/zero-pressure mesh), the POD motion mode basis set can be expressed as:

$$\Psi_0 = [\psi_1 \cdots \psi_M]. \quad (2.26)$$

Thus, the deformation can be expressed as

$$\mathbf{u} \approx \Psi_0 \mathbf{a}. \quad (2.27)$$

where $\mathbf{a}$ denotes the modal coefficients predicted by the network.

This reduces the output from thousands of nodal displacement values to only 10–20 modal coefficients, which simplifies learning and can improve optimisation stability when training high-dimensional models. The reported results showed the proposed model was about $30\times$ faster than the reference FEA model (including training time), while achieving average errors of -3% in EF and 7% in PS. However, although this approach is promising for clinical use, a key limitation is its dependence on the POD basis. To generalise across diverse anatomies, the basis must be built from a sufficiently large and representative dataset so that the dominant motion patterns of clinically encountered ventricular shapes are covered.

Among these approaches, PINNs can be viewed as a promising alternative to FEA because they enforce the governing equations without requiring dense displacement labels, and their mechanics can be evaluated in an FEA-like manner using element shape functions. However, robust convergence remains challenging for complex cardiac PDEs. While the variants above introduce practical strategies to improve optimisation stability, they still have notable limitations, such as reliance on reduced-order motion bases and demonstrations on idealised geometries. Further development is therefore needed before PINN-based methods can meet routine clinical requirements.

2.4.3 Graph-based Surrogates

In addition to early MLP and PINN-based architectures, significant progress has been made in the development of graph-based data-driven surrogates. GNNs operate directly on graph data, where

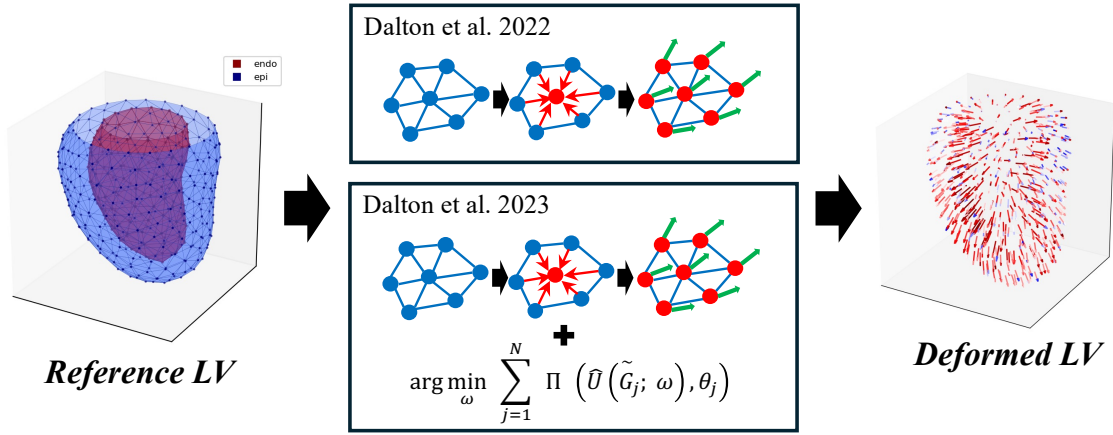


Figure 2.4: **Two representative Graph-based surrogates for LV mechanics:** [18] employed an encoder–GNN processor–decoder architecture to model LV deformation, [19] extended this framework to a PI-GNN architecture, incorporating biomechanical constraints during training.

nodes represent spatial points and edges encode their interactions, enabling structured information propagation to learn complex system behaviour. A general GNN can be represented as:

$$\mathbf{h}_i^{(k+1)} = \psi^{(k)} \left(\mathbf{h}_i^{(k)}, \sum_{j \in \mathcal{N}(i)} \phi^{(k)} \left(\mathbf{h}_i^{(k)}, \mathbf{h}_j^{(k)}, \mathbf{e}_{ij} \right) \right). \quad (2.28)$$

where i indexes a mesh node, $\mathcal{N}(i)$ denotes its set of neighbouring nodes defined by mesh connectivity, $\mathbf{h}_i^{(k)}$ is the node feature (encoding geometry and/or physical state) at message-passing layer k , $\mathbf{e}_{ij}$ is an optional edge feature (e.g., relative position or edge length) for the connection (i, j) , $\phi^{(k)}$ is the learnable message function, $\psi^{(k)}$ is the learnable update function, and the summation aggregates messages from all neighbours to update $\mathbf{h}_i^{(k+1)}$.

The representative GNN-based studies on ventricular FEA have largely been shaped by two works from Dalton *et al.*, which outline the recent SOTA development in this area (Fig. 2.4).

In 2022, Dalton *et al.* proposed a GNN-based framework to emulate LV FEA simulations during diastole [18]. The model takes as input the reference mesh geometry together with simulation parameters (e.g., cavity pressure), and outputs the nodal displacement field. The network follows an encoder–processor–decoder architecture. The encoder embeds geometry and loading into latent node features, the processor performs multiple rounds of message passing to capture spatial mechanical interactions, and the decoder maps the latent representation back to physical outputs (displacements). Their results showed good agreement with FEA displacement fields while achiev-

ing orders-of-magnitude speed-up, demonstrating an effective forward surrogate under fixed mesh topology and prescribed loading.

Building on this emulator, Dalton et al. subsequently introduced a physics-informed GNN (PI-GNN) to improve physical consistency in ventricular FEA emulation [19]. Compared with the purely data-driven GNN, the PI-GNN augments the training objective with additional physics-based constraints, encouraging predictions that better respect the underlying biomechanical behaviour of the ventricular wall. This extension retains the core encoder–processor–decoder architecture, but regularises learning via physics-guided supervision. The PI-GNN demonstrated improved stability and reduced non-physical artefacts in the predicted displacement fields, particularly under unseen loading conditions, while preserving the computational efficiency of the emulator.

Unsupervised Cross-Modal Motion Tracking over the Cardiac Cycle

S. Mu et al., “Imc-pinn-fe: A physics-informed neural network for patient-specific left ventricular finite element modeling with image motion consistency and biomechanical parameter estimation,” *arXiv preprint arXiv:2506.20696*, 2025 (under review)

Contents

3.1 Introduction	29
3.1.1 Contributions	31
3.2 Methods	31
3.2.1 B-spline Pairwise Registration	32
3.2.2 Loss function design: Fourier regularization	33
3.2.3 Connection Methods	34
3.3 Experiments and Results	35
3.3.1 Datasets	35
3.3.2 Evaluation factors	37
3.4 Experiments and Results	37
3.4.1 2D Evaluation	37
3.4.2 3D Evaluation	42
3.5 Discussion	42
3.6 Conclusion	44

Cardiac motion estimation through image registration of medical images is essential for assessing myocardial strains and stroke volume to assess cardiac function during diseases, and to support image-based biomechanics simulations. DL has proven to be a powerful approach for image registration, being faster and more accurate than conventional numerical methods. However, Deep Learning Image Registration (DLIR) algorithms are typically trained for a specific image

type and modality (MRI, CT or ECHO), and the training cannot be transferred to another image type/modality. The efficacy is often constrained by the difficulty in getting enough images of a particular type/modality. Therefore, it is necessary to develop common models, thereby reducing the specific issues of any single modality, such as resolution limitations or excessive noise. In this chapter, we propose the Elastix Fourier Deep Learning (*EFDL*) framework for such a cross-type/modality unsupervised DLIR training. Unlike image-based methods, *EFDL* does not use traditional convolution kernel to extract and process visual information, but introduces a B-spline grid to divide the entire image training into point displacement training. Our results showed that *EFDL* has good performance in 2D+t/3D+t motion tracking for ECHO and MRI cardiography, exceeding some popular DLIR algorithms and close to computationally expensive, highly regularized gold standard non-DL algorithm from [41]. Also, to verify the cross-modality, we tested multiple combinations of training and test datasets with different modalities. It has been demonstrated that training with a combination of datasets improves the performance of *EFDL*, but degrades other tested algorithms, demonstrating the ability of *EFDL* utilize to cross-type/modal training.

3.1 Introduction

4D cardiac cine imaging, commonly used in MRI and ECHO, is a non-invasive technique that provides time-resolved observations of cardiac motion throughout the cardiac cycle [94]. This imaging modality produces continuous sequences of 2D images that capture the dynamic behaviour of the heart over time. Such data provide valuable insights into cardiac function and ventricular dynamics. However, interpreting these 2D image sequences to understand the underlying three-dimensional motion of cardiac structures remains challenging, even for experienced clinicians. Consequently, accurate algorithms for tracking cardiac motion from 2D or 3D image sequences are essential. In clinical conditions such as cardiomyopathy [95] and myocardial infarction [96], reliable motion tracking enables physicians to analyse pathological changes by examining dynamic cardiac morphology and biomechanical indicators such as myocardial strain and stroke volume [97].

Over the past decades, a wide range of algorithms have been developed to estimate cardiac motion. These methods can generally be divided into traditional approaches and data-driven approaches. Traditional techniques such as speckle tracking [98], optical flow tracking [31], B-spline Fourier tracking [41], and surface-based tracking [99] have demonstrated good personalised pairwise registration performance in specific medical imaging modalities, including MRI, CT, and ECHO. However, these optimisation-based methods typically involve lengthy processing times, which limits

their clinical applicability. In recent years, data-driven algorithms based on DL have gained increasing attention due to their rapid inference speed. These approaches can be broadly categorised into supervised methods, which require labelled training data, and unsupervised methods, which operate without explicit motion labels. Because obtaining large labelled motion datasets is difficult in medical imaging, unsupervised methods are becoming increasingly practical for real-world applications.

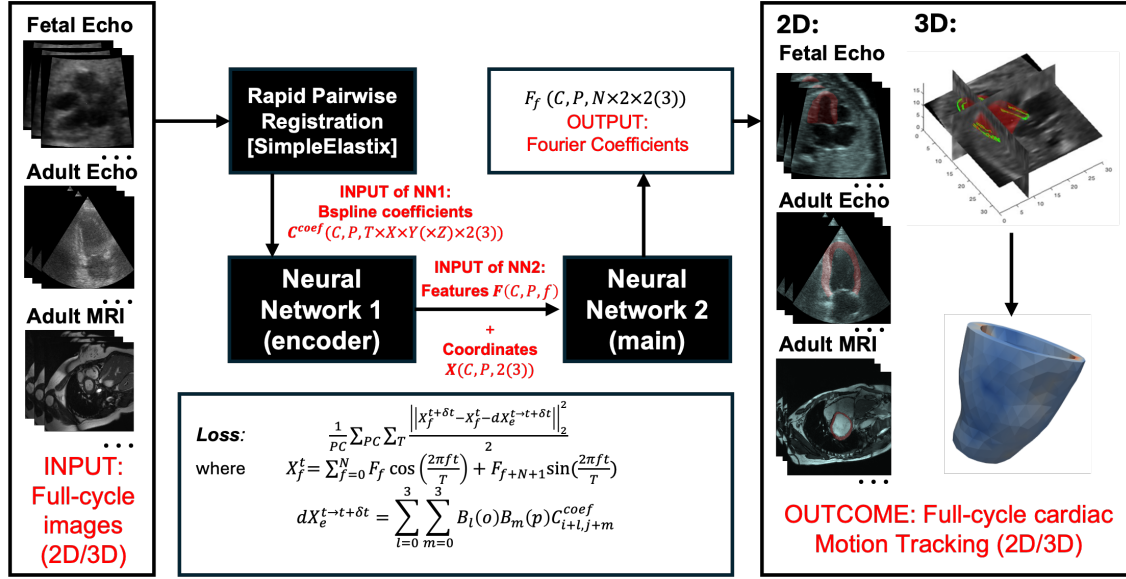


Figure 3.1: **Schematic of the *EFDL* Framework:** The input consists of 3D/4D full-cycle cardiac images (ECHO/MRI). Initially, these images are processed through a rapid elastix pairwise registration module, which generates B-spline coefficients delineating the approximate transformations between each pair of neighbouring images. Subsequently, Neural Network 1, functioning as an encoder, reduces the dimensionality of the features. Neural Network 2, the primary network, executes Fourier regularization. The output, comprising Fourier coefficients, facilitates the back-computation of displacements across all time points. This enables the tracking of cyclic movements in the myocardium, with the resultant tracking outcomes displayed.

Despite these advances, current unsupervised motion tracking algorithms still present several limitations. First, most algorithms focus on pairwise registration, which primarily enforces spatial consistency between image pairs while neglecting temporal consistency across the entire cardiac cycle. Although pairwise registration results may appear satisfactory, accumulated errors can lead to degraded performance when these pairwise transformations are combined to obtain sequence-level tracking. Second, existing unsupervised motion tracking approaches have mainly demonstrated strong performance in imaging modalities such as MRI and CT, whereas their performance in ultrasound imaging remains less reliable, particularly for 3D motion tracking [100]. In addition, there is currently no single algorithm capable of robustly tracking cardiac motion across multiple

imaging modalities simultaneously. Third, to support pathological analysis, incorporating physics-based PDE constraints into unsupervised motion tracking to recover biomechanically meaningful quantities such as stress and stiffness has become an active research direction. However, most CNN- and U-Net-based tracking methods represent motion as voxel-wise discrete displacement fields, which do not naturally align with the continuous or element-wise function spaces required for enforcing strong or weak PDE constraints [22, 101].

To address these challenges, this work proposes Elastix Fourier Deep Learning (*EFDL*), an FCNN-based framework for unsupervised motion tracking across multiple imaging modalities. The *EFDL* framework consists of two main components: B-spline registration and Fourier regularisation. The B-spline registration provides a rapid non-deep-learning solution for coarse pairwise alignment, while the Fourier regularisation module employs a deep learning model to enforce temporal consistency across the cardiac cycle. This personalised coarse registration step extracts global motion features that facilitate the subsequent deep learning model in reducing appearance differences between image modalities and improving cross-modality tracking performance.

3.1.1 Contributions

Our main contributions are threefold:

1. **Unsupervised motion tracking with temporal consistency.** We develop an unsupervised spatio-temporal FCNN-based motion tracking framework that estimates coherent motion trajectories across frames without requiring ground-truth displacement supervision.
2. **Cross-modality compatibility.** We combine B-spline based global registration with the proposed learning-based tracker to improve robustness across imaging modalities, and demonstrate that the same framework is applicable to both MRI and ECHO.
3. **2D pre-trained and 3D patient-specific deployment.** We provide a 2D pre-trained model evaluated on both MRI and ECHO datasets for data-driven performance analysis, and further present a patient-specific model for 3D tracking.

3.2 Methods

The architecture of the *EFDL* framework (Fig 3.1) is divided into two principal components: rapid pairwise registration using B-spline transformations and Fourier regularization. To integrate these

components, three connection methods, termed 'ADD', 'MODE', and 'MIX', are employed and compared. Subsequent sections detail the technical aspects of the B-spline pairwise registration, Fourier regularization and various connection approaches.

3

3.2.1 B-spline Pairwise Registration

Theoretically, the model could be adapted to any pairwise registration algorithm. In this project, we employed a cubic B-spline deformation algorithm, using the Python package SimpleElastix [102, 103] due to its universality and high accuracy in MRI and ultrasound image registration. SimpleElastix was an established medical image registration library that used adaptive stochastic gradient descent to derive an optimal B-spline transformation from randomly distributed samples within an image. It used Mean Square Error (MSE) to measure similarity between the B-spline interpolation results and the ground truth. In addition, the algorithm incorporated a bending energy term to regularize the deformation, with equal relative weighting across components. In our implementation, we used the default setting in which B-spline control points were spaced at regular intervals of 10 units of the maximum voxel dimension. Although effective, this default setting was not necessarily optimal, suggesting that future work could make the B-grid spacing adaptive. Finally, in this step we chose Eulerian registration (neighbour registration) rather than Lagrangian registration (first-frame registration), since Eulerian registration typically involved smaller deformations and produced lower errors.

After completing the B-spline Elastix pairwise registration, SimpleElastix output B-spline coefficients as $\mathbf{C}^{\text{coef}}(C, P, T \times X \times Y(\times Z) \times \text{dim})$, where C represented the number of training/testing cases, P denoted the number of sampled coordinates across the image (distinct from the exact number of pixels, since points were typically skipped at regular intervals to ensure smoothness), T indicated the number of time points per cycle, $X \times Y(\times Z)$ was the coefficient grid between neighbouring time points (with Z included for 3D), and dim referred to the dimension of the domain (2 for 2D, 3 for 3D). Using $\mathbf{C}^{\text{coef}}(X, Y(\times Z))$, the displacements at the next time point ($t + \delta t$) could be calculated as,

$$2D : \quad d\mathbf{X}_{t \rightarrow t+\delta t} = \sum_{l=0}^3 \sum_{m=0}^3 B_l(o) B_m(p) \mathbf{C}_{i+l, j+m}^{\text{coef}} \quad (3.1)$$

$$3D : \quad d\mathbf{X}_{t \rightarrow t+\delta t} = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(o) B_m(p) B_n(q) \mathbf{C}_{i+l, j+m, k+n}^{\text{coef}} \quad (3.2)$$

where $\mathbf{X}_{t \rightarrow t+\delta t}$ denotes the deformation vector field from time point t to $t + \delta t$ (with $\delta t = 1$ in our case), and B_n refers to the four cubic B-spline basis functions [104]. The terms o, p (and q for 3D) were defined as $(o, p, q)^T = \mathbf{X}_t - \lfloor \mathbf{X}_t \rfloor$, where $\mathbf{X}_t$ was the displacement at time point t within the B-spline reference domain. The terms i, j (and k for 3D) were defined as $(i, j, k)^T = \lfloor \mathbf{X}_t \rfloor - 1$. The term $\mathbf{C}_{i+l, j+m, k+n}^{\text{coef}}$ referred to the B-spline coefficients at the control point with lattice indices i, j (and k for 3D).

3.2.2 Loss function design: Fourier regularization

After obtaining the external Elastix deformation information, the model then trained Neural Network 1 and Neural Network 2 simultaneously, as illustrated in Fig. 3.1. The input to Neural Network 2 comprised a concatenation of the B-spline latent features and image coordinates, where the coordinates were flattened into tensors of size $(C, P, 2)$ for 2D cases or $(C, P, 3)$ for 3D cases, with P denoting the number of sampled coordinates. In this project, we used images of size 512×512 , which corresponded to an input coordinate tensor of size $(C, 262144, 2)$ for 2D cases (and $(C, 262144 \times 40, 3)$ for 3D cases when a temporal dimension of 40 frames was included). To encourage smoothness, the coordinates were subsampled on a regular grid. Specifically, we set the sampling gap to 8, resulting in an effective coordinate size of $(C, 64 \times 64, 2)$ for 2D and $(C, 64 \times 64 \times 40, 3)$ for 3D.

The output of Neural Network 2 was a set of Fourier coefficients $\mathbf{F}_f$, from which the displacement at time t was reconstructed as

$$\mathbf{X}^t = \sum_{f=0}^N \left(\mathbf{F}_f \cos\left(\frac{2\pi ft}{T}\right) + \mathbf{F}_{f+N+1} \sin\left(\frac{2\pi ft}{T}\right) \right), \quad (3.3)$$

where $\mathbf{X}^t$ denotes the displacement at time t , T is the period of the cardiac cycle, and N is the number of retained Fourier modes. Each coefficient satisfies $\mathbf{F}_f \in \mathbb{R}^{\text{dim}}$, where dim is the number of Cartesian spatial dimensions (2 for 2D and 3 for 3D), and f indexes the frequency mode.

In the main network, the loss function was defined as the accumulated discrepancy between the Fourier-based deformation and the pairwise Elastix deformation across time pairs:

$$\mathcal{L} = \frac{1}{CPT} \sum_{c=1}^C \sum_{p=1}^P \sum_{t=1}^T \frac{\|\mathbf{X}_{c,p}^{t+\delta t} - \mathbf{X}_{c,p}^t - d\mathbf{X}_{e,c,p}^{t \rightarrow t+\delta t}\|_2^2}{2}, \quad (3.4)$$

where C is the number of training/testing cases, P is the number of sampled coordinates, and T

is the number of registered time pairs. The terms $\mathbf{X}_{c,p}^{t+\delta t}$ and $\mathbf{X}_{c,p}^t$ denote the Fourier-reconstructed displacements at two time points for coordinate p in case c . The term $d\mathbf{X}_{e,c,p}^{t \rightarrow t+\delta t}$ denotes the pairwise Elastix displacement from t to $t+\delta t$, computed by inserting $\mathbf{X}_{c,p}^t$ into Eq. (3.1) or Eq. (3.2). Notably, this consistency regularization was not restricted to temporally adjacent pairs and could be applied to arbitrary registered time pairs. In this study, we used neighbour-to-neighbour registration with uniform weights, although non-uniform weighting could be adopted to reflect different practical requirements.

3.2.3 Connection Methods

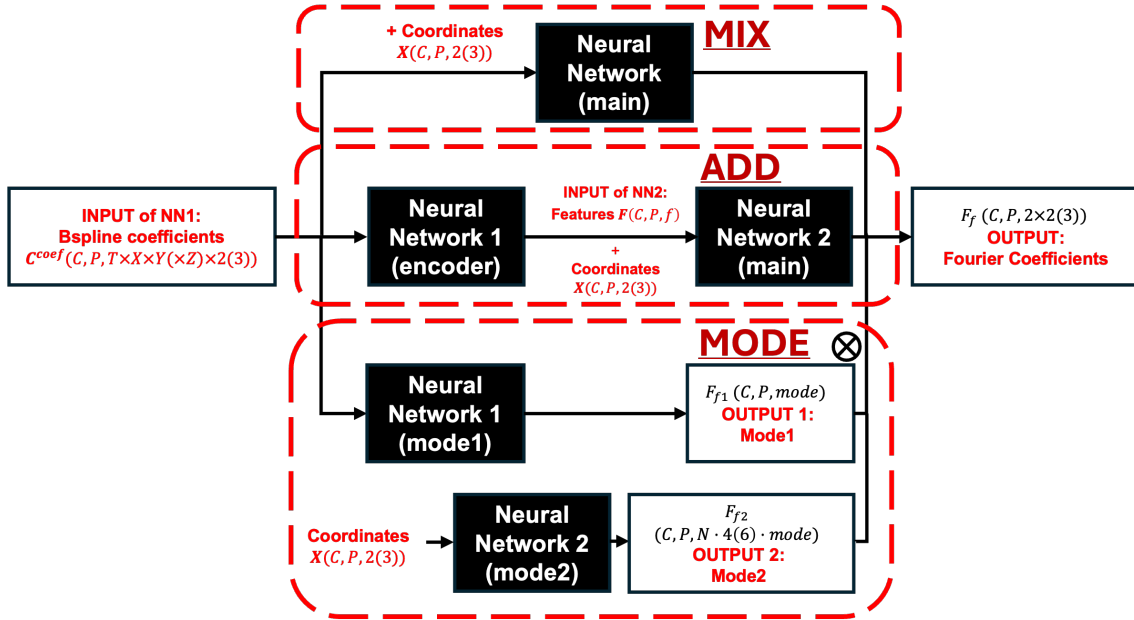


Figure 3.2: Connection approaches: MIX, ADD, and MODE.

To accurately predict myocardium motion, our model incorporated two types of information: the coordinates of the image prior to movement and the B-spline coefficients that represented the movement trajectory of a specific myocardial point in the B-grid. To effectively integrate these two sets of information, we explored three common methods of connection. These methodology enabled the neural network to learn from both spatial features and displacement fields.

MIX Connection. An intuitive method for combining these two types of information involved directly inputting them into a FCNN. This method (top part in Fig. 3.2) was straightforward to implement. However, it did not adequately consider the differing importance of the two types of information, primarily due to the disproportionate quantities of data each represented. This

imbalance in the data weights could adversely affect the model's performance.

ADD Connection. This connection method incorporated two FCNNs. The first FCNN processed the B-spline coefficients into reduced-order features. These processed features were then input into the second FCNN, together with the coordinates of the static image. The output from the second FCNN consisted of Fourier coefficients that described the motion throughout the entire cardiac cycle. These coefficients were subsequently utilised within the loss function to update both neural networks, enhancing the model's predictive accuracy and efficiency.

MODE Connection. In our MODE connection approach, we also employed two FCNNs to process and integrate the data. The first FCNN analysed B-spline coefficients and output a predefined number of features, designated as *MODE1*, with the shape $(C, P, mode)$. The second FCNN processed the spatial coordinates to reduce their dimensions, aligning the output dimensions to match those of the first network, designated as *MODE2* shaped as $(C, P, 4(6) \cdot N \cdot mode)$. The outputs of both FCNNs were then subjected to matrix-dot multiplication to derive the Fourier coefficients, with the calculation $((4(6) \cdot N, mode) \cdot (mode, 1) = 4(6) \cdot N)$.

3.3 Experiments and Results

3.3.1 Datasets

The 2D+t and 3D+t datasets used in this study are described in the following sections.

3.3.1.1 2D+t datasets

The 2D *EFDL* framework was tested on four datasets with different modalities, including the public adult cardiac MRI dataset ACDC [105], the public healthy adult echocardiography dataset CAMUS [106], and our private fetal ECHO datasets (healthy and aortic stenosis).

ACDC (adult 2D MRI) [105]. The ACDC dataset was acquired from clinical exams at the University Hospital of Dijon using two MRI scanners with different magnetic field strengths (1.5 T and 3.0 T). In this project, we selected exams that covered a complete cardiac cycle from the 150 available cases and extracted 2D slices from each *.nii.gz* volume. The resulting 2D MRI dataset

comprised 951 training cases and 538 testing cases. Manual segmentations of myocardium and the LV were provided as ground truth.

3

CAMUS (adult 2D ECHO) [106]. CAMUS is a 2D ECHO dataset acquired at the University Hospital of St Etienne (France) using GE Vivid E95 scanners. It consisted of 500 healthy adult subjects with both apical two-chamber (A2C) and apical four-chamber (A4C) views. Following the official split, 400 cases were used for training and 100 for testing. Manual segmentations of myocardium and LV were provided.

Private fetal 2D ECHO [41]. Our private fetal dataset consisted of 4D ECHO acquisitions, from which 2D ECHO videos were obtained at multiple slices. The dataset included 97 videos from 15 healthy fetuses and 81 videos from 9 diseased fetuses (aortic stenosis). Data were acquired using the Spatio-Temporal Image Correlation (STIC) mode with a GE Voluson 730 ultrasound system equipped with an RAB 4-8L transducer (GE Healthcare Inc., Chicago, IL, USA) at National University Hospital, Singapore [41]. For the 2D *EFDL* setup, 70 cases were used for training and 27 cases for testing. The dataset was manually labelled (LV and myocardium) at ES and ED frames, and labels for intermediate frames were propagated using a gold-standard cardiac motion estimation algorithm [41].

Since *EFDL* was trained in an unsupervised manner, segmentation labels were used only for evaluation and not for optimisation. Specifically, the ACDC and CAMUS myocardium masks were provided as ground truth, while the fetal ECHO datasets were manually segmented using LazySnap [107]. To improve robustness, we applied standard data augmentation, including flipping, cropping, and rescaling, resulting in an effective training set approximately four times larger than the original set.

3.3.1.2 3D+t datasets

For 3D+t evaluation, we used two multi-slice datasets introduced above: the ACDC public adult cardiac MRI dataset [105] and our private fetal ECHO dataset [41, 42]. Specifically, we included 50 adult MRI cases from ACDC and 5 personalised fetal ECHO cases (3 healthy and 2 diseased subjects with aortic stenosis). Both datasets provide 2D slices that can be stacked to reconstruct a 3D volume at each time frame, enabling a 3D+t evaluation over one cardiac cycle. Each case consisted of 20–40 temporal frames.

3.3.2 Evaluation factors

The experiments were evaluated using the provided ground-truth motion vectors. We report the DSC, MSE, and HD [108], defined in Eqs. (3.5)–(3.7).

$$DSC = \frac{1}{N} \sum_{i=1}^N \frac{2|M_{p,i} \cap M_{t,i}|}{|M_{p,i}| + |M_{t,i}|} \quad (3.5)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (I_{p,i} - I_{t,i})^2 \quad (3.6)$$

$$HD = \frac{1}{2N} \sum_{i=1}^N \left(\frac{1}{|M_{p,i}|} \sum_{x_1 \in M_{p,i}} \min_{x_2 \in M_{t,i}} d(x_1, x_2) + \frac{1}{|M_{t,i}|} \sum_{x_1 \in M_{t,i}} \min_{x_2 \in M_{p,i}} d(x_1, x_2) \right) \quad (3.7)$$

where N is the number of sampling points, $M_{p,i}$ represents binary value of the predicted warped myocardium mask, $M_{t,i}$ is the binary value of the true myocardium mask, $I_{p,i}$ is the pixel value of the predicted warped raw image, $I_{t,i}$ is the pixel value of the true raw image, $d(x_1, x_2)$ denotes the shortest distance between mask x_1 and mask x_2 . The MSE estimates the local pixel-level similarity, the DSC score quantifies the amount of similarity between the predicted and true myocardium masks, and the HD measures edge roughness and irregularity [109].

All *EFDL* experiments were conducted using the NVIDIA Modulus framework (PyTorch v1.4.0). The model was trained using a MLP architecture consisting of 4 layers, each containing 1024 neurons. All *EFDL* models were optimised using the Adam optimizer with a learning rate of 0.001 and a decay rate of 0.95.

3.4 Experiments and Results

3.4.1 2D Evaluation

3.4.1.1 Effect of B-grid Size, Number of Fourier Coefficients, and Connection Method

We first conducted 2D ablation experiments to examine the impact of the B-grid size, the number of Fourier coefficients N , and the feature connection method. The results are summarised in

No.	Data	N	B-grid	Conn.	DSC $\uparrow$	HD (mm) $\downarrow$	MSE $\downarrow$	Time (s) $\downarrow$
Fourier coefficient number N								
1	Adult ECHO [A2C]	4	14×14	ADD	0.931 ± 0.032	3.34 ± 1.07	0.0017 ± 0.0006	51
2	Adult ECHO [A2C]	3	14×14	ADD	0.887 ± 0.052	4.10 ± 1.23	0.0021 ± 0.0013	45
3	Adult ECHO [A2C]	8	14×14	ADD	0.931 ± 0.039	3.49 ± 1.20	0.0016 ± 0.0006	82
B-grid size								
4	Adult ECHO [A2C]	4	7×7	ADD	0.927 ± 0.022	3.99 ± 1.56	0.0017 ± 0.0009	43
5	Adult ECHO [A2C]	4	21×21	ADD	0.929 ± 0.018	3.21 ± 1.49	0.0015 ± 0.0008	96
6	Adult MRI	4	7×7	ADD	0.798 ± 0.082	6.81 ± 1.62	0.0590 ± 0.0134	54
7	Adult MRI	4	14×14	ADD	0.870 ± 0.078	5.22 ± 1.61	0.0049 ± 0.0030	59
8	Adult MRI	4	21×21	ADD	0.881 ± 0.092	4.48 ± 2.82	0.0052 ± 0.0021	103
9	Fetal ECHO [A4C]	4	7×7	ADD	0.802 ± 0.068	5.98 ± 2.30	0.0586 ± 0.0179	39
10	Fetal ECHO [A4C]	4	14×14	ADD	0.890 ± 0.049	4.03 ± 1.08	0.0027 ± 0.0009	44
11	Fetal ECHO [A4C]	4	21×21	ADD	0.886 ± 0.051	4.10 ± 1.97	0.0024 ± 0.0005	81
Connection method								
12	Adult ECHO [A2C]	4	14×14	MIX	0.901 ± 0.017	3.83 ± 2.52	0.0019 ± 0.0010	48
13	Adult ECHO [A2C]	4	14×14	MODE	0.930 ± 0.021	3.51 ± 1.95	0.0017 ± 0.0009	52

Table 3.1: **Ablation study on N , B-grid size, and connection method. Row 1 is the reference setting.**

Table 3.1. The MSE was computed on normalised motion fields and is therefore unitless.

Rows 1–3 evaluate different values of N . Increasing N from 3 to 4 improved the DSC by about 5 percentage points, while further increasing N to 8 yielded only marginal changes. The MSE followed a similar trend, and the lowest HD was achieved at $N = 4$. Considering both accuracy and runtime, we selected $N = 4$ for the remaining experiments.

Rows 1 and 4–11 compare B-grid sizes across three datasets. A coarse grid can underfit and miss local motion details, while a dense grid can be more sensitive to noise and increases optimisation cost. Overall, 14×14 and 21×21 achieved comparable performance, but 21×21 required nearly twice the runtime. We therefore used a 14×14 B-grid to maintain a consistent and efficient setting across datasets.

Rows 1 and 12–13 compare connection methods. The MIX variant performed worse across all metrics, suggesting that it reduced the contribution of the coordinate input. The ADD and MODE variants were comparable, with ADD slightly better overall. We therefore adopted ADD in the subsequent sections.

3.4.1.2 Quantitative Comparison Under Single-Modal and Mixed Training

In this section, we compared 2D *EFDL* with two representative registration baselines: (i) BSF [41, 42], an extension of SimpleElastix that incorporates temporal Fourier-based regularisation

No.	Train	Test	Algorithm	DSC $\uparrow$	HD (mm) $\downarrow$	MSE $\downarrow$	Time (s) $\downarrow$
Single-modal training and testing							
1	AE	AE	EFDL	0.931 ± 0.032	3.34 ± 1.07	0.0017 ± 0.0006	51
			VoxelMorph	0.864 ± 0.051	4.33 ± 1.34	0.0026 ± 0.0009	1
			BSF	0.923 ± 0.019	3.31 ± 1.01	0.0019 ± 0.0008	1198
2	FE	FE	EFDL	0.890 ± 0.049	4.03 ± 1.08	0.0027 ± 0.0008	49
			VoxelMorph	0.864 ± 0.051	4.81 ± 1.34	0.0035 ± 0.0019	1
			BSF	0.910 ± 0.017	3.98 ± 1.04	0.0029 ± 0.0011	1125
3	AM	AM	EFDL	0.870 ± 0.078	5.22 ± 2.61	0.0049 ± 0.0030	59
			VoxelMorph	0.832 ± 0.073	5.85 ± 3.74	0.0384 ± 0.0102	1
			BSF	0.904 ± 0.061	4.12 ± 2.93	0.0041 ± 0.0019	1189
Cross-dataset generalisation							
4	AE	FE	EFDL	0.847 ± 0.059	4.96 ± 1.12	0.0042 ± 0.0028	49
			VoxelMorph	0.762 ± 0.098	6.35 ± 4.35	0.0177 ± 0.0096	1
5	AM	FE	EFDL	0.819 ± 0.062	5.08 ± 1.42	0.0052 ± 0.0031	49
			VoxelMorph	0.598 ± 0.213	7.98 ± 4.65	0.0483 ± 0.0137	1
6	FE	AM	EFDL	0.810 ± 0.071	5.92 ± 3.82	0.0463 ± 0.0103	59
			VoxelMorph	0.699 ± 0.166	6.87 ± 4.98	0.0596 ± 0.0291	1
Mixed-dataset training							
7	AE+FE	FE	EFDL	0.903 ± 0.016	4.03 ± 1.24	0.0027 ± 0.0004	49
			VoxelMorph	0.869 ± 0.046	4.51 ± 1.28	0.0031 ± 0.0008	1
8	AE+FE	AE	EFDL	0.937 ± 0.029	3.13 ± 1.07	0.0016 ± 0.0007	51
			VoxelMorph	0.896 ± 0.037	3.98 ± 1.11	0.0020 ± 0.0010	1
9	AE+FE+AM	AM	EFDL	0.881 ± 0.063	4.65 ± 1.91	0.0031 ± 0.0014	59
			VoxelMorph	0.786 ± 0.092	6.31 ± 3.94	0.0426 ± 0.0188	1
10	AE+FE+AM	FE	EFDL	0.901 ± 0.030	3.89 ± 1.06	0.0025 ± 0.0006	49
			VoxelMorph	0.803 ± 0.084	5.11 ± 2.22	0.0049 ± 0.0034	1
11	AE+FE+AM	AE	EFDL	0.934 ± 0.021	3.28 ± 1.00	0.0015 ± 0.0003	51
			VoxelMorph	0.784 ± 0.122	6.78 ± 4.01	0.0103 ± 0.0072	1
Legend: AE, Adult ECHO [A2C]; FE, Fetal ECHO [A4C]; AM, Adult MRI.							

Table 3.2: **Quantitative motion tracking results under single-modal, cross-dataset, and mixed-dataset training.**

to enforce cyclic motion, and (ii) VoxelMorph [101], a widely adopted DL-based image registration method for pairwise registration. Table 3.2 summarises the results under three settings: single-modal evaluation, cross-dataset generalisation, and mixed-dataset training. For each case, evaluation factors were averaged over the full cardiac cycle.

Single-modal baselines: Single-modal evaluation trains and tests on the same dataset, using different subjects for training and testing. On Adult ECHO, *EFDL* achieved the highest DSC and the lowest MSE, while BSF obtained the lowest HD. On Fetal ECHO, BSF slightly improved DSC

and HD, whereas *EFDL* yielded the lowest MSE. On Adult MRI, BSF was the most accurate across DSC, HD, and MSE. VoxelMorph consistently ran fastest but showed lower accuracy across all three datasets.

Although BSF tended to be marginally more accurate, *EFDL* provided a substantially faster learning-based alternative. In our implementation, *EFDL* was about 20–24 $\times$ faster than BSF, with only a small loss in accuracy (within about 4 percentage points in DSC across the single-modal settings). The qualitative motion tracking results remained visually comparable.

Cross-dataset generalisation: Cross-dataset evaluation trains and tests on different datasets. Under this setting, we compared the learning-based methods only. *EFDL* consistently outperformed VoxelMorph across all metrics, achieving higher DSC with lower HD and MSE. The largest gap was observed for Adult MRI $\rightarrow$ Fetal ECHO, where VoxelMorph degraded markedly, indicating higher sensitivity to modality-dependent appearance changes. This trend suggests that the B-spline parameterisation in *EFDL*, which constrains the motion to a smooth, low-dimensional deformation space, reduces reliance on modality-specific appearance statistics and improves robustness under distribution shifts.

Mixed-dataset training: Mixed-dataset training constructs the training set by combining multiple datasets, while evaluation is performed on the held-out test split of one dataset within the mixture. When the mixture involved the same modality (Adult ECHO + Fetal ECHO, No. 7–8), both *EFDL* and VoxelMorph improved across all metrics relative to their single-modal counterparts, indicating that additional ECHO data and diversity benefited generalisation. In contrast, when the mixture included different modalities (MRI + ECHO, No. 9–11), *EFDL* achieved consistent gains across all three test domains, whereas VoxelMorph degraded. This behaviour further suggests that the B-spline coefficient representation in *EFDL* is more robust to heterogeneous training distributions. By comparison, the displacement fields predicted by the dense convolutional neural network are more flexible and may be more susceptible to the appearance features of specific modalities. Therefore, simply mixing MRI and ECHO may reduce performance.

3.4.1.3 Cycle-level Temporal Consistency

Figure 3.3 provides representative qualitative examples together with the DSC evolution over an entire cardiac cycle. Visually, the propagated contours remained well aligned with the myocardial

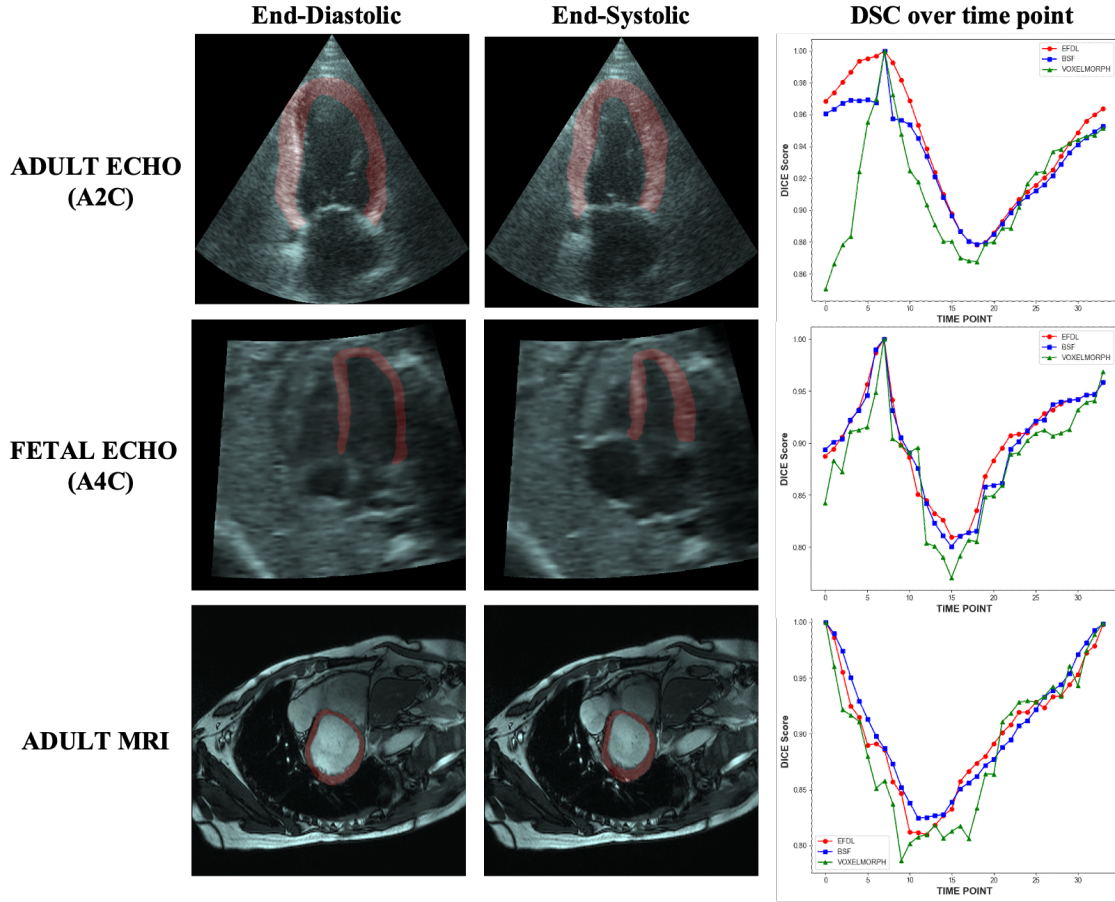


Figure 3.3: **Representative tracking visualisations (left) and per-frame DSC trajectories (right) for Adult ECHO, Fetal ECHO, and Adult MRI over one cardiac cycle.** The reference frame used for deformation has DSC = 1 in each plot.

boundaries at two key physiological time points, ED and ES. In each DSC-over-time plot, the single point reaching DSC = 1 corresponds to the reference frame, and all other frames were obtained by deforming this reference towards the target time points. Across the three examples, *EFDL* and *BSF* consistently achieved higher and more stable DSC trajectories than *VoxelMorph*. The performance gap became more evident as the temporal distance from the reference frame increased, where *VoxelMorph* showed a clearer drop in overlap. This behaviour is consistent with the absence of an explicit periodic temporal constraint in *VoxelMorph*, whereas the Fourier-based cyclic regularisation in *EFDL* and *BSF* better preserved long-range consistency throughout the cycle.

3.4.2 3D Evaluation

We further evaluated *EFDL* for patient-specific 3D+t motion tracking on multi-slice Adult MRI and Fetal ECHO data. Following the 2D study, we compared *EFDL* against Elastix [102] (a classical traditional registration approach without temporal regularization) and Co-attention STN (a SOTA DL 3D+t image registration approach [100]), together with the previously introduced BSF and VoxelMorph baselines. Table 3.3 summarises the results, time is the training hours.

For the Adult MRI dataset, *EFDL* achieved the highest DSC (0.829 ± 0.026) and the lowest HD (4.71 ± 0.17), indicating strong registration accuracy. Co-attention STN and BSF delivered comparable performance (within 0.014 in DSC and 0.19 mm in HD). However, BSF required at least $2.4\times$ longer runtime. Co-attention STN can provide fast inference after training, but it must be retrained when transferred to a different image type or modality. Such supervised training is time-consuming and requires a large labelled dataset, which is often difficult to obtain. In addition, Co-attention STN required image down-sampling due to GPU memory constraints, which can limit the achievable accuracy. By contrast, *EFDL* does not require large-scale supervised pre-training, but it is optimised per subject and typically completes within 1.6 hours per case.

For the Fetal ECHO dataset, Co-attention STN achieved the best DSC and HD. *EFDL* achieved the next best performance, with comparable accuracy (DSC 0.802 ± 0.030 , HD 4.99 ± 0.45).

Overall, Co-attention STN and *EFDL* are suitable for different practical regimes. Co-attention STN is attractive when sufficient labelled data and compute are available for modality-specific training, enabling near real-time inference. In scenarios with limited labelled data or when cross-modality deployment is required, *EFDL* provides a more practical balance between accuracy and per-case runtime. Unless otherwise stated, subsequent experiments used *EFDL* as the registration method.

3.5 Discussion

Across both 2D and 3D evaluations, *EFDL* demonstrated a consistent and robust performance profile, highlighting its suitability as a general-purpose cardiac motion tracking framework. In 2D+t experiments, *EFDL* achieved accuracy comparable to highly regularised optimisation-based method (BSF) while substantially reducing runtime, and it consistently outperformed conventional DL baselines (VoxelMorph) under cross-dataset and mixed-dataset settings. In 3D+t evaluations,

Method	Data	DSC $\uparrow$	HD (mm) $\downarrow$	Time (h) $\downarrow$
No registration	Adult MRI	0.681 ± 0.109	6.61 ± 1.12	–
	Fetal ECHO	0.623 ± 0.120	6.83 ± 1.61	–
Elastix [102]	Adult MRI	0.736 ± 0.112	5.97 ± 1.33	0.4
	Fetal ECHO	0.729 ± 0.108	6.01 ± 1.05	0.4
BSF [41, 42]	Adult MRI	0.815 ± 0.030	4.86 ± 0.34	3.9
	Fetal ECHO	0.803 ± 0.008	4.94 ± 0.26	4.0
VoxelMorph [101]	Adult MRI	0.766 ± 0.067	5.85 ± 1.10	8.1
	Fetal ECHO	0.703 ± 0.051	6.11 ± 0.96	8.3
Co-attention STN [100]	Adult MRI	0.817 ± 0.031	4.90 ± 0.22	7.4
	Fetal ECHO	0.804 ± 0.012	4.92 ± 0.19	7.2
<i>EFDL</i>	Adult MRI	0.829 ± 0.026	4.71 ± 0.17	1.6
	Fetal ECHO	0.802 ± 0.030	4.99 ± 0.45	1.7

Table 3.3: **Quantitative comparison of 3D+t cardiac motion estimation methods on Adult MRI and Fetal ECHO.** Time is training time

EFDL maintained this favourable balance, achieving SOTA accuracy on Adult MRI and competitive performance on Fetal ECHO, with per-case runtimes that were practical for subject-specific analysis. Importantly, the same framework and parameterisation were applied across modalities and dimensionalities without modification, indicating strong scalability from 2D slice-based tracking to full 3D+t motion estimation.

A key strength of *EFDL* lies in its unsupervised formulation and its explicit enforcement of temporal consistency. Instead of relying on ground-truth displacement supervision, *EFDL* estimates pairwise motion using a B-spline deformable registration framework, while the Fourier-based temporal regularisation suppresses high-frequency temporal variations to enforce smooth and periodic motion across the cardiac cycle. The B-spline representation constrains deformation to a smooth, low-dimensional space, while the Fourier model provides an explicit cyclic prior and couples all time points within a cardiac cycle. This temporal coupling improves long-range consistency, particularly for frames far from the reference, and reduces sensitivity to modality-specific appearance variations. Therefore, *EFDL* generalises well across imaging modalities and continues to benefit from heterogeneous training data, whereas dense CNN-based displacement predictors can more easily overfit modality-dependent intensity cues. Overall, these results indicate that *EFDL* offers a practical and robust alternative for full-cycle cardiac motion tracking when labelled data are limited and temporal coherence is required.

This study has several limitations. The current 3D+t implementation of *EFDL* follows a patient-specific optimisation paradigm rather than a fully pre-trained model, due to constraints in hardware resources and the limited availability of large-scale 3D+t datasets. While this ap-

proach is substantially faster than classical optimisation-based methods, it does not yet achieve near real-time performance (within one minute as observed in the 2D+t setting). With improved computational resources and access to larger, more diverse 3D+t datasets, future work could extend *EFDL* towards a fully pre-trained deployment to further reduce inference time. Hence, when real-time performance is critical and the target domain is restricted to a single modality with stable acquisition conditions, we recommend Co-attention STN as a strong practical alternative for 3D+t registration.

We did not conduct extensive sensitivity analyses in this chapter. This choice was motivated by the close methodological relationship between *EFDL* and BSF, as both rely on the same B-spline deformation model combined with cyclic Fourier temporal regularisation. The primary contribution of *EFDL* lies in accelerating the solution of this optimisation problem through a learning-based solver, rather than altering the underlying deformation formulation. Detailed sensitivity analyses of the BSF framework have been thoroughly investigated in prior studies [41, 42], and are therefore not repeated here.

3.6 Conclusion

In this chapter, we presented *EFDL*, an unsupervised full-cycle cardiac motion tracking framework that enforces temporal consistency through Fourier regularisation in a smooth B-spline deformation space. The key strength of *EFDL* is that it provides a robust cross-modal solution rather than serving only as a strong single-dataset tracker. It achieved competitive single-modal accuracy, substantially outperformed the conventional DL baseline under cross-dataset and mixed-dataset settings, and remained effective in both 2D+t and 3D+t applications across MRI and ECHO. In particular, it showed clear robustness to heterogeneous training distributions and offered a favourable trade-off between registration quality and runtime. However, *EFDL* is not without limitations. Its best-case single-modal accuracy can still be slightly below that of BSF, its 3D+t fetal ECHO performance was marginally below that of Co-attention STN, and its current 3D+t implementation still relies on patient-specific optimisation rather than fully pre-trained real-time inference. Therefore, the main contribution of *EFDL* is not that it dominates every baseline on every metric, but that it offers a more general, temporally coherent, and practically deployable framework for full-cycle motion tracking when labelled data are limited and cross-modality applicability is required.

Reduced-Order Physics-Informed Neural Networks for Left Ventricular Mechanics over the Cardiac Cycle

S. Mu et al., “Imc-pinn-fe: A physics-informed neural network for patient-specific left ventricular finite element modeling with image motion consistency and biomechanical parameter estimation,” *arXiv preprint arXiv:2506.20696*, 2025 (under review)

Contents

4.1 Introduction	46
4.1.1 Contributions	48
4.2 Methods	49
4.2.1 Myocardial Stiffness and Maximal Tension Estimator	50
4.2.2 Volume-Constrained PINN-FEA Solver	55
4.2.3 Neural Network Architecture	57
4.3 Experiments and Results	60
4.3.1 Performance of $C_{\text{stiffness}}$ and T_{max} Estimator	60
4.3.2 Performance of PINN-FEA solver	62
4.4 Discussion	67
4.4.1 Sensitivity and Error Analysis	68
4.4.2 Limitations and Future Work	69
4.5 Conclusion	70

Elucidating the biomechanical behaviour of the myocardium is important for understanding cardiac physiology. This cannot be directly inferred from clinical imaging, but requires FEA simulations. However, conventional FEA simulations are computationally expensive and often do not accurately reproduce clinically observed cardiac motions. In this chapter, we propose the *IMC-PINN-FE*, a PINN framework that integrates imaged motion consistency (IMC) with FEA for patient-specific simulations of LV biomechanics. Cardiac motion is first estimated from MRI or ECHO using either a pre-trained attention-based network (Co-attention STN [100]) or an unsupervised cyclic-regularised network (*EFDL* in Chapter 3), followed by extraction of motion modes. *IMC-PINN-FE* then estimates myocardial stiffness and active tension by fitting clinical pressure measurements within seconds. Based on these parameters, *IMC-PINN-FE* performs FEA modelling throughout the cardiac cycle and achieves an overall ~ 75 -fold speed-up over conventional FEA simulations. Through motion constraints, its outputs of cardiac displacements across the cardiac cycle match imaged motions much more accurately, improving DSC from 0.849 to 0.927 compared to conventional volume-constrained FEA methods. Concurrently, it preserves a realistic P-V relationship that is very similar to that of conventional FEA. *IMC-PINN-FE* improves on existing PINN-FEA models by enabling back-computation of myocardial properties and achieving better imaged motion fidelity. This strategy of using cardiac motions of a single individual to reconstruct shape modes avoids the need for a large dataset, and enables better patient-specificity. *IMC-PINN-FE* thus presents a robust approach for rapid, patient-specific, and image-consistent cardiac biomechanical modelling.

4.1 Introduction

Computation of the biomechanical behaviour of the heart is useful for evaluating cardiovascular diseases and can facilitate the development of cardiac digital twins for optimised and personalised medicine. Cardiac biomechanics is traditionally computed through image-based FEA simulations. This has previously been performed for both healthy and pathological cardiac conditions [2, 3], providing details of myocardial deformation, stress distribution, and contractile function [6], which offered physiological insights into diseases and intervention scenarios [57].

However, current FEA-based approaches face significant challenges in clinical applications. Traditional cardiac FEA methods rely on time-consuming simulations, making real-time, patient-specific modelling impractical [110]. Additionally, most FEA models assume population-averaged

myocardial properties, such as stiffness and active tension, and are thus not fully patient-specific [70]. Moreover, existing FEA methods often struggle to faithfully replicate the complex, time-dependent myocardial deformations observed in MRI and ultrasound data [111]. While iterative tuning of the FEA model to back-compute myocardial properties and to better match clinically measured motion is possible [112], it is both time and resource intensive. Therefore, FEA models have not yet been widely adopted in clinical practice. This highlights the need for an efficient, patient-specific, and physiologically consistent alternative, one that is fast enough to be clinically adoptable and that integrates image-based motion tracking with physics-informed biomechanical constraints.

In recent years, DL based FEA surrogates have developed rapidly, providing substantial acceleration of cardiac FEA simulations [18–20, 24, 113]. A number of data-driven LV FEA surrogate models have been proposed, among which the more advanced approaches are graph-based architectures that operate directly on unstructured meshes and can accommodate variable mesh topology while achieving orders-of-magnitude speed-ups for passive ventricular mechanics [18–20]. However, these pre-trained graph surrogates typically rely on large offline FEA datasets, assume fixed or population-averaged material properties, and are largely confined to the passive diastolic phase, without modelling active contraction.

As an alternative to such pre-trained surrogates, PINN offer another route that combines deep learning with biomechanical physics constraints [88]. Unlike pre-trained models, PINN-based formulations are inherently patient-specific and therefore more flexible, allowing myocardial material parameters to be varied and re-estimated across different cases or clinical scenarios. Several PINN-FEA approaches have recently been proposed for modelling LV biomechanics. Motiwale et al. developed a PINN-FEA framework that predicts the nodal displacements of an idealised 3D LV mesh from pressure and active tension inputs, guided by the governing equations in their weak form [21]. Buoso et al., in contrast, incorporated a pre-computed statistical shape model to encode cardiac geometry and motion, enabling a similar input-output mapping across a variety of realistic cardiac anatomies [22]. Their use of shape encoding significantly reduced computational costs. However, these methods do not enforce an adherence to imaged myocardial motion, which limits their adaptability to patient-specific variations. Moreover, they lack a mechanism for inverse estimation of myocardial stiffness and active contractile tension.

To address these challenges, we propose Image Motion Consistent PINN-FEA (*IMC-PINN-FE*), a novel framework that integrates patient-specific motion tracking into a reduced-order PINN-FEA model. *IMC-PINN-FE* encodes cardiac motion using modes extracted from cine-imaging data of

an individual subject rather than from a large dataset of diverse individuals. This helps to retain the computational efficiency of reduced-order modelling while eliminating the requirement for an extensive dataset and enhances the patient-specificity of modelled motions. Furthermore, the framework is constrained to match observed LV volumes from clinical images, and it incorporates inverse estimation of myocardial stiffness and peak active tension prior to performing detailed biomechanical simulations.

4

To provide the motion input required by *IMC-PINN-FE*, we used *EFDL* (described in chapter 3), which enforces cyclic consistency in pairwise image registration and can be applied to both cine MRI and 4D echocardiography. As an alternative, we also suggest a previously reported cross-attention-based deep learning network [100] for scenarios with sufficient training data and near-real-time requirements. Together with standard 3D LV segmentation and mesh-reconstruction pipelines [107, 114–116], these components provide the anatomical geometry and 4D motion fields required for personalised *IMC-PINN-FE* modelling.

4.1.1 Contributions

Our main contributions are threefold:

1. **Motion-Consistent PINN-FEA.** We propose the *IMC-PINN-FE*, a framework for rapid tissue biomechanics modelling of the cardiac ventricle that enforces consistency with both physics-based governing equations and 4D image-derived cardiac motion, thereby improving the physiological accuracy of biomechanical assessments. This represents a departure from prior approaches that primarily focused on enforcing physics constraints. Motion consistency is achieved through the use of patient-specific cardiac motion modes and constraints based on imaged ventricular volumes.
2. **High Computational Efficiency.** *IMC-PINN-FE* leverages image-derived motion modes for reduced-order modelling, enabling a reduction in computational time by approximately $75\times$ compared to traditional FEA methods, while maintaining fidelity to patient-specific motion. This efficiency makes the approach feasible for real-time or near real time clinical applications. Further, our use of cine-images of a single patient rather than a large multi-individual dataset for statistical shape modelling [22] is less cumbersome and more subject-specific.
3. **Efficient Physics-Informed Parameter Estimation.** *IMC-PINN-FE* performs fast, un-

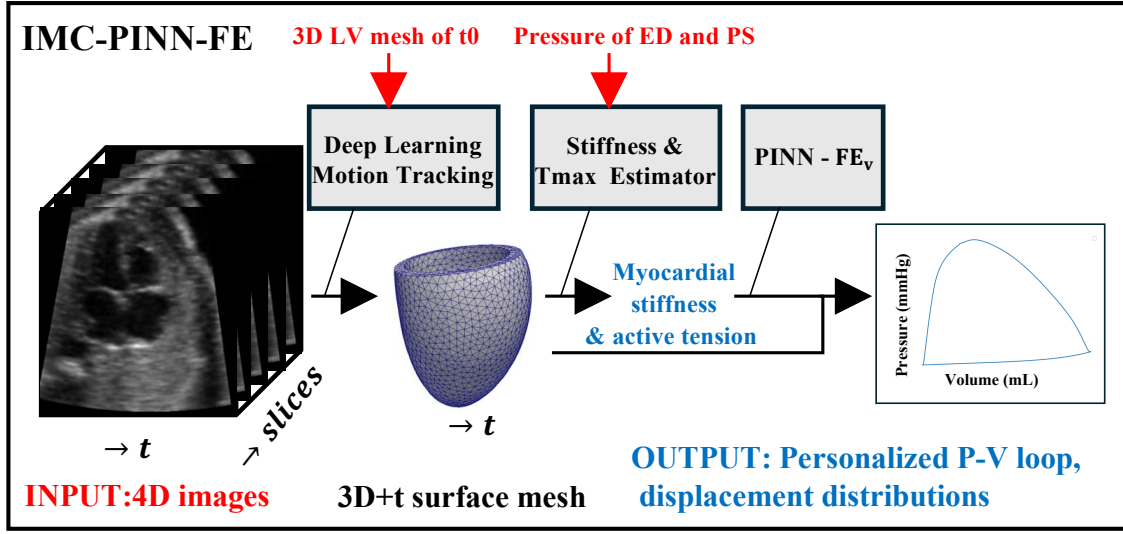


Figure 4.1: **The overall framework of *IMC-PINN-FE*.** Inputs to the framework include 4D (3D+time) imaging data (e.g., ECHO or MRI) combined with LV reconstruction obtained via DL and physiological inputs, including PS and ED LV pressures. Outputs include estimates of patient-specific LV myocardial stiffness and maximum active tension, and LV FEA simulation results across the cardiac cycle.

supervised estimation of myocardial stiffness and peak active tension, which enhances personalisation and supports accurate patient-specific biomechanical modelling.

4.2 Methods

The framework of *IMC-PINN-FE*, as illustrated in Fig. 4.1, follows a three-step pipeline that integrates deep learning and physics-informed principles:

1. **DL Cardiac Motion Tracking (Detailed in Chapter 3):** This step aims to extract smooth and cyclic cardiac motion from 4D medical images. Two distinct approaches are presented: (a) our 3D+t *EFDL* method introduced in Chapter 3 (3D schematic in Fig. 4.2), and (b) a previously reported cross-attention deep learning network [100]. The former does not require pre-training on large-scale datasets but is more time-intensive than the latter during inference.
2. **Back-Computing Myocardial Stiffness and Active Tension:** Using the estimated motion and inputted LV pressure information, a PINN network predicts myocardial stiffness and maximum active tension, based on FEA governing equations.
3. **PINN-FE solver:** This final step uses the same solver as in (2) to perform FEA modelling

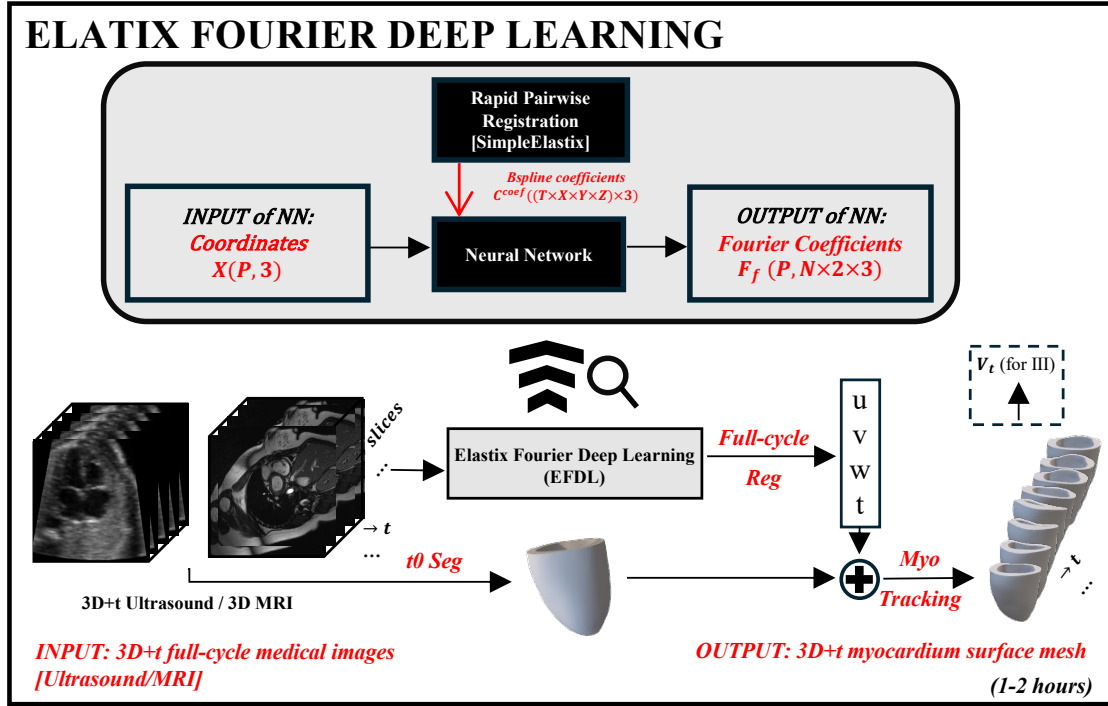


Figure 4.2: **Workflow of the 3D EFDL in IMC-PINN-FE:** The input consists of 4D full-cycle cardiac images (ECHO/MRI). Initially, these images undergo processing through a rapid Elastix pairwise registration module, which computes B-spline coefficients representing the approximate transformations between consecutive image pairs for coarse registration. Next, a neural network applies Fourier regularization using coordinate-based inputs. The output, comprising Fourier coefficients, enables the reconstruction of displacements across all time points. When combined with manual segmentation of the initial frame t_0 using LazySnap [107] for coarse delineation and Geomagic (<https://oqton.com/geomagic-designx>) for 3D refinement, this framework enables the tracking of cyclic myocardial motion.

for the whole cardiac cycle, where shape mode coefficients describing displacements are obtained under the constraints of the governing equations and imaged LV volumes guidance.

4.2.1 Myocardial Stiffness and Maximal Tension Estimator

The myocardial stiffness and maximum tension estimator was implemented using PINN, which first estimated the coefficients of the stiffness model, followed by the maximum contractile tension. The overall procedure was illustrated in Fig. 5.1 II and consisted of the following steps.

1. **3D Cardiac Mesh Reconstruction:** The 3D surface mesh of the LV myocardium was first reconstructed through manual segmentation of a single reference frame using LazySnap [107] for coarse delineation, followed by surface refinement in Geomagic <https://oqton.com/geomagic-designx>. However, this can similarly be achieved with DL segmentation and reconstruction algorithms [114–116]. The resulting surface mesh was then converted into

a tetrahedral FEA mesh using Gmsh [117]. This reference FEA mesh was subsequently propagated to all other time points using motion fields estimated by the *EFDL* framework. In addition, at each time point, the cavity volume was computed by sealing the endocardial surface and performing volumetric integration on the surface mesh. The resulting volume served as a reference for the next step in performing the volume constraint in the PINN-FEA solver.

2. Load-Free State Definition: The reference zero-pressure unloaded geometry was assumed to correspond to the cardiac state at 1/3 into the diastolic duration, as this phase typically exhibited the lowest blood cavity pressure. This assumption was imperfect, but consistent with prior studies [79, 118, 119]. At this reference time point, the myocardial helix angle was assumed to vary linearly from the epicardium to the endocardium (values given in appendix 8.1). The following parameters were precomputed at the load-free state:

- Fiber (f_{unloaded}), sheet (s_{unloaded}), and sheet-normal (n_{unloaded}) orientations. (methods given in Chapter 2.3)
- Nodal endocardial area (A_{unloaded}) and myocardial volume (V_{unloaded}) were computed by distributing the area of each triangular surface element and the volume of each tetrahedral element to their associated nodes via weighted summation [22].
- Mesh labels ($Label_{\text{unloaded}}$) [interior: 0, endocardium: 1, epicardium: 2, base: 3] were assigned to each node.
- The spatial derivative operator at each node (∇_0) was applied, as explained in Section 4.2.1.1.

3. Motion Analysis via POD: A previous PINN-FEA model [22] encoded cardiac motion using shape modes derived from POD applied to cardiac geometries reconstructed from multiple subjects at a fixed phase of the cardiac cycle. However, this approach captured inter-subject anatomical variability rather than intra-subject dynamic deformations. Here, we applied POD to the series of cardiac shapes across the cardiac cycle from a single subject. This enabled the extraction of subject-specific dominant motion modes, which were then used for compact and physiologically meaningful motion encoding. The resulting motion modes were used as a reduced-order representation of displacement fields in the PINN-FEA solver, which improved numerical stability and accelerated convergence.

4. Stiffness and Maximal Tension Estimation: Two similar neural networks were trained to estimate myocardial stiffness ($C_{\text{stiffness}}$) and maximum active tension (T_{max}), using motion

mode amplitudes at ED and PS as inputs. The physical constraints incorporated into the loss functions are identical to the FEA equations presented in Section 2.3 (all predefined parameters in these equations, based on the literature, are shown in appendix 8.2), and the two networks are described in detail in Section 4.2.1.3.

4.2.1.1 Reference Gradient Operator ∇_0

The reference displacement gradient $\nabla_0 \mathbf{u}$ appears in Eq. (2.14) through the term $\mathbf{P} : \nabla_0 \delta \mathbf{u}$ and is required to form $\mathbf{F} = \mathbf{I} + \nabla_0 \mathbf{u}$. In our implementation, $\nabla_0 \mathbf{u}$ is approximated via a precomputed discrete gradient operator ($\mathbf{S}$) constructed from the reference mesh, avoiding AD of spatial coordinates inside the network.

We followed Mancinella et al. [120] to calculate the discrete gradient operator. To estimate the gradient of a scalar function $f(x)$ defined over tetrahedral (TET4) meshes (nodes of tetrahedral meshes in our case), for each element, the gradient was first evaluated at its centroid $\nabla f(\mathbf{x}_c) = (\nabla f(\mathbf{x}_c)_x, \nabla f(\mathbf{x}_c)_y, \nabla f(\mathbf{x}_c)_z)$ by solving a local linear system constructed from the coordinates and function values of the element's four vertices v_j, v_i, v_k, v_l as,

$$\begin{bmatrix} \mathbf{v}_j - \mathbf{v}_i \\ \mathbf{v}_k - \mathbf{v}_i \\ \mathbf{v}_l - \mathbf{v}_i \end{bmatrix} \begin{bmatrix} \nabla f(\mathbf{x}_c)_x \\ \nabla f(\mathbf{x}_c)_y \\ \nabla f(\mathbf{x}_c)_z \end{bmatrix} = \begin{bmatrix} f(\mathbf{x}_j) - f(\mathbf{x}_i) \\ f(\mathbf{x}_k) - f(\mathbf{x}_i) \\ f(\mathbf{x}_l) - f(\mathbf{x}_i) \end{bmatrix}. \quad (4.1)$$

The resulting element-wise gradients were subsequently aggregated to the nodes via a volume-weighted average over all neighbouring tetrahedra sharing each node, which defines the discrete nodal gradient operator $\mathbf{S}$ (∇f in Eq. 4.1). Through this technique, these local derivative matrices could be precomputed and assembled into a global matrix operator outside the network, allowing efficient evaluation of nodal gradients via matrix–vector multiplication when using the network. To avoid numerical instability, elements involving apex nodes were excluded from the assembly due to their small element size and disproportionate sensitivity to local deformation fluctuations near the apex, which could amplify numerical noise and destabilise the reduced-order decomposition [22].

4.2.1.2 Proper Orthogonal Decomposition (POD)

The use of POD shape modes in *IMC-PINN-FE* is inspired by previous work [121–124]. Before applying POD, all meshes underwent spatial alignment for consistency. Specifically, each mesh

was translated such that the centre of the basal plane aligned with the origin of the Cartesian coordinate system. The mesh was then rotated to orient the basal plane to be normal to the positive z -axis.

The mesh size for an individual's LV was selected to contain between 2,000 and 3,000 nodes, balancing computational efficiency with low discretisation error. A total of 20 to 40 temporal frames (depending on data acquisition protocols) across the cardiac cycle were included in the POD computation. Displacement was defined as the deviation of each node from its position in the identified load-free state. These displacement fields were concatenated across time into a high-dimensional data matrix, and POD was applied via Singular Value Decomposition (SVD) to extract a set of orthonormal spatial basis functions [22, 121]. The resulting basis functions were ranked in descending order of their contribution to the total variance, and the first 10 dominant modes were retained to represent the subject-specific deformation subspace. The displacement at any time point, denoted as u , was approximated as a linear combination of these basis functions:

$$u = \Psi \vec{A}_{n,t} \quad (4.2)$$

where $\Psi = [\Psi_1, \dots, \Psi_n]$ represents the set of POD basis functions (in this project, we use the first 20 bases), and the term $\vec{A}_{n,t}$ represents the amplitudes corresponding to each basis function.

4.2.1.3 Estimation Framework of $C_{\text{stiffness}}$ and T_{max}

Stiffness Estimation: The stiffness estimator network took as input the amplitudes of the POD modes extracted from the ED frame, and output the myocardial stiffness coefficient $C_{\text{stiffness}}$, while the POD motion bases and the LV cavity pressure at ED were treated as predefined parameters. The stiffness coefficient represents the global stiffness scaling factor, and the optimisation was conceptually akin to inferring this factor by simulating the diastolic pressure loading process that deformed the ventricle from the load-free state to the ED configuration. For this estimation task, only the passive component of the strain energy was considered ($W = W_{\text{passive}}$), and active contraction was assumed to be absent.

Clinically, direct measurement of LV ED pressure (Left Ventricular End-Diastolic Pressure (LVEDP) P_{ED} (mmHg)) is challenging, as it typically requires invasive catheter-based procedures. However, empirical equations derived from prior clinical studies enabled estimation of P_{ED} using non-invasive imaging data. For cardiac MRI, we applied empirical formulations provided by

Burkhoff et al.[125]:

$$P_{ED} = 0.015 \left(e^{0.06(V_{ED}-V_0)} - 1 \right)$$

where V_{ED} is the ED volume obtained from MRI, and V_0 is the theoretical unloaded volume of the LV at zero pressure.

For echocardiography, P_{ED} was estimated based on the ratio between mitral inflow Doppler velocity (E) and mitral annular early diastolic Doppler velocity (E'), as proposed by [126]:

$$P_{ED} = 0.85 \times \frac{E'}{E} + 4.4 \quad (4.3)$$

The computed stiffness coefficient was used for maximum tension estimation in the next network.

Maximal Tension Estimation: The maximal tension estimator took as inputs the amplitudes of the POD modes extracted from the PS frame and output the maximal active tension T_{max} , while the POD motion basis, stiffness coefficient, and LV cavity pressure at PS were treated as predefined parameters. Here, we assumed that the timing of the PS frame was known. Practically, it could be obtained from tonometry or photoplethysmography combined with Electrocardiogram (ECG). At this stage, both active and passive components of the strain energy were considered. The LV pressure at PS could be non-invasively measured from the upper arm via a standard pressure cuff.

4.2.1.4 Datasets and Evaluation

To evaluate the accuracy of the back-computed parameters ($C_{stiffness}$ and T_{max}), we performed FEA simulations using predefined values of $C_{stiffness}$ and T_{max} , which served as ground truths. From these simulations, a limited number of time frames were extracted and used as cardiac motion estimation inputs in the estimators (as illustrated in Fig. 4.3). The estimated $C_{stiffness}$ and T_{max} values were then compared against the ground truth to assess the accuracy of the estimator.

In total, 8 LV FEA simulations were performed. 2 were based on healthy adult human image data from the *Cardiac Atlas Project* <https://www.cardiacatlas.org/>, 2 were based on healthy fetal echocardiography images from our private dataset [41, 42], and 4 were based on porcine echocardiography images from a previous study [126], including 2 healthy pig hearts and 2 with LV hypertrophy induced by surgical aortic constriction. These simulations therefore span a range

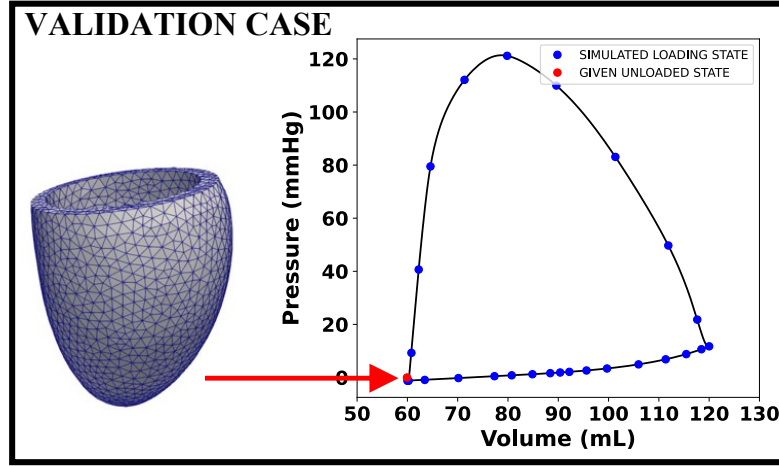


Figure 4.3: **Schematic of validation case generation.** A limited number of states (blue points) are selected from the FEA simulation, which is performed using a given unloaded surface mesh (red points), to serve as tracking results.

of LV sizes and include both healthy and pathological conditions.

In addition to the FEA-based validation, we further tested our estimator on 2 cases of real clinical images, using *EFDL* for motion tracking: 1 adult case MRI from the *Cardiac Atlas Project* and 1 fetal case from our private dataset [41]. Since the ground truth myocardial stiffness and active tension are not known, we adopted values from literature [70, 127, 128] as ground truths. These are not ideal ground truths, but applying the framework to these real cases served to demonstrate its practical feasibility.

4.2.2 Volume-Constrained PINN-FEA Solver

The personalised PINN-FEA solver was employed to model LV biomechanics over the entire cardiac cycle (Fig 5.1 III), utilizing the encoded motion modes and the personalised stiffness and active tension parameters estimated in the previous step. The solver incorporated an FCNN to predict the LV pressure P_t and mesh displacements u_t at each time point t within the cardiac cycle, using the elapsed time since ED (t) and the corresponding LV cavity volume V_t as inputs. Both t and V_t were uniformly sampled across the cardiac duration, based on coarse frames derived from the tracking results in Section 2.1. In our implementation, the cardiac duration was assumed to be 800 ms for adults [70] and 400 ms for fetuses [128]. These durations can be adapted in future applications to accommodate patient-specific timing. The node displacement outputs were in the form of amplitudes of the motion basis $A_{n,t}$, and could be decoded as per Eq. 4.2. Subsequently, stress could also be computed via Eq. 2.11. *IMC-PINN-FE* was volume-constrained in the sense

that it directly adopted the LV volume over time obtained from the images, and only computed the corresponding pressure. It did not interface with a lumped parameter model to determine the LV volume. However, such a model could easily be incorporated in future.

4.2.2.1 Loss function for PINN-FEA solver

The loss function of this network consisted of 6 components.

1. The physics-governing equation, Eq. 4.4.
2. The volume control equation, which ensured that LV volume was consistent with images, Eq. 4.5.
3. The ED pressure control equation, which enforced the ED pressure at the ED frame, Eq. 4.6.
4. The PS pressure control equation, which enforced the PS pressure at the PS frame, Eq. 4.6.
5. The diastolic phase slope constraint, which ensured a continuous increase in the P-V loop slope during diastole to aid convergence, Eq. 4.7.
6. The systolic phase slope constraint, which ensured a continuous increase in the P-V loop slope to aid convergence, Eq. 4.7.

$$loss_{energy} = \frac{1}{T} \sum_T \left(\int_{\Omega_{myo}} W dV - \int_{\Omega_{endo}} P \cdot u dA \right)^2 \quad (4.4)$$

$$loss_{volume} = \frac{1}{T} \sum_t (V_{predict} - V_{target})^2 \quad (4.5)$$

$$loss_{ED/PS} = (P_{ED/PS,predict} - P_{ED/PS,target})^2 \quad (4.6)$$

$$loss_{dia/sys} = \sum_{T_{dia/sys}} \max \left(0, \frac{\Delta P_{t_{dia/sys}}}{\Delta V_{t_{dia/sys}}} - \frac{\Delta P_{t_{dia/sys}+1}}{\Delta V_{t_{dia/sys}+1}} \right) \quad (4.7)$$

where T represents the total number of time points, $V_{predict}$ and V_{target} denote the predicted and target myocardial volumes, respectively, $P_{ED/PS,predict}$ and $P_{ED/PS,target}$ represent the predicted and target LV pressures at the ED/PS frame, and $\frac{\Delta P_{t_{dia/sys}}}{\Delta V_{t_{dia/sys}}}$ denotes the slope of P-V loop in the

diastolic or systolic phase.

Hence, the overall loss function was

$$\begin{aligned} loss = & w_1 loss_{energy} + w_2 loss_{volume} + \\ & w_3 loss_{ED} + w_4 loss_{PS} + \\ & w_5 loss_{dia\ trend} + w_6 loss_{sys\ trend} \end{aligned} \quad (4.8)$$

where w_i represents the weight of each component, as defined in appendix 8.3.

The comparison group was the set of volume-constrained LV FEA simulations, generated using the same equations as in Section 2.3.

4.2.2.2 Datasets and Evaluation

To evaluate the performance of the volume-constrained PINN-FEA solver, we conducted the complete procedure of *IMC-PINN-FE* in Fig. 5.1 to model cardiac biomechanics across the entire cardiac cycle. We compared *IMC-PINN-FE* with the traditional FEA simulation and the previously proposed PINN-FEA method by Buoso et al. [22]. We assessed whether the simulations maintained a closeness to the imaged motions, and whether they produced realistic P-V loops.

This evaluation was conducted on 6 human cases introduced in section 4.3.1, comprising 3 adult cardiac MRI cases and 3 fetal echocardiography cases. The ground truth for imaged cardiac motion was defined as 4D LV mesh reconstructions propagated across the cardiac cycle using *EFDL* due to its highest accuracy without requiring a large training dataset.

The P-V loop was evaluated using the RE at key pressure points, including EDP and PSP, together with the coefficient of determination (R^2) for the overall P-V relation, as defined in appendix 8.4.

4.2.3 Neural Network Architecture

For *IMC-PINN-FE* I, all *EFDL* experiments were conducted using the NVIDIA Modulus framework [129] (PyTorch v1.4.0). The model was trained using a MLP architecture consisting of 4

layers, each containing 1024 neurons. All *EFDL* models were optimised using the Adam optimizer with a learning rate of 0.001 and a decay rate of 0.95.

For *IMC-PINN-FE* II, both estimators employed an MLP with 3 hidden layers, each containing 32 neurons, and were trained to converge by minimizing Eq. 2.11 to zero. Both estimators were optimised using the AdamW optimizer with a learning rate of 0.001.

For *IMC-PINN-FE* III, the PINN-FEA neural network was configured with 4 hidden layers, each containing 128 neurons, and was trained to converge by minimizing Eq. 2.14. The framework was optimised using the AdamW optimizer with a learning rate of 0.0001 and a decay rate of 0.8.

All models were executed on NVIDIA GeForce RTX 3090 Ti PCIe 3.0 GPUs with 24GB GDDR6X memory.

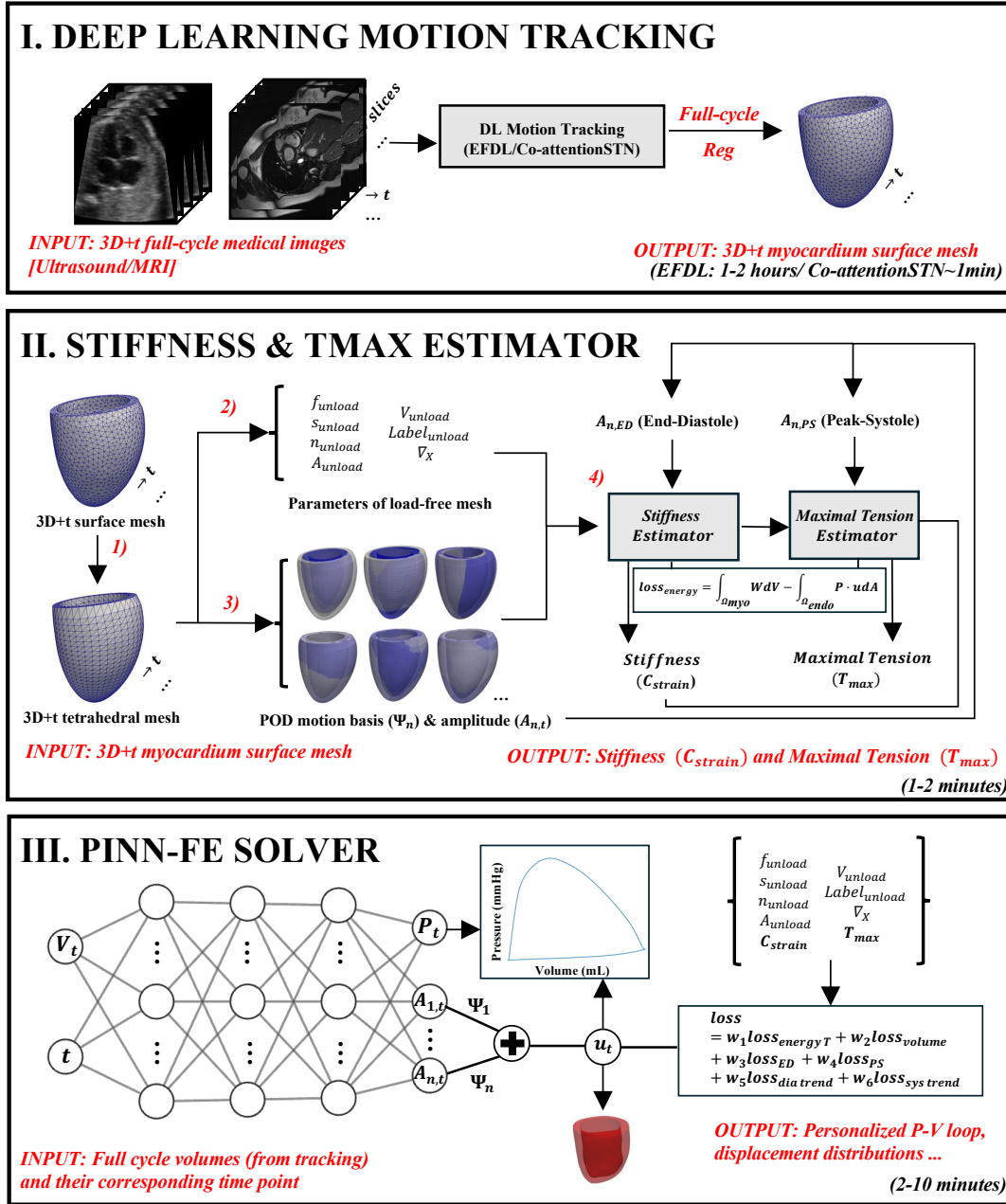


Figure 4.4: **Workflow of the IMC-PINN-FE framework.** **I. Deep Learning Motion Tracking:** 4D full-cycle cardiac images (Echo/MRI) are used to generate myocardial motion meshes via either *EFDL* (our proposed method, no pre-training required) or Co-attention STN (previous SOTA method, pre-training required). **II. Myocardial Stiffness and Maximal Tension Estimator:** Starting from the 3D+t surface mesh, the pipeline includes: (1) tetrahedral meshing via Gmsh; (2) preprocessing to define helix angles and the load-free reference state; (3) POD to extract motion bases and amplitudes across frames; and (4) personalized estimation of myocardial stiffness ($C_{stiffness}$) and maximal active tension (T_{max}) using ED and PS amplitudes. **III. Volume-Constrained PINN-FEA Solver:** Inputs include myocardial volumes and their time points (from Section 2.1), sampled uniformly from the cardiac cycle starting at ED. A PINN with MLP architecture predicts LV pressure P_t and motion amplitudes $A_{n,t}$, from which displacement fields u_t are reconstructed via Eq. 4.2, enabling personalized P-V loop and biomechanical analysis.

4.3 Experiments and Results

4.3.1 Performance of $C_{\text{stiffness}}$ and T_{max} Estimator

Table 4.1: **Physiological parameters of the 2 clinical image-tracking cases and 8 FEA-simulated validation cases.** EDV and ESV are derived from 3D image reconstructions. EDP and PSP for image tracking cases follow physiological values reported in prior clinical and experimental studies [130–132].

Source	Case	EDV (mL)	ESV (mL)	EDP (mmHg)	PSP (mmHg)
Image tracking (Real)	Adult-1	120.88	71.31	11.0 [130, 131]	120.0 [130, 131]
	Fetal-1	0.41	0.22	5.0 [132]	27.0 [132]
FEA tracking (Validation)	Adult-2	121.12	60.01	11.0	120.0
	Adult-3	135.36	54.23	11.0	120.0
	Fetal-2	0.39	0.20	2.0	21.8
	Fetal-3	0.63	0.28	2.0	32.3
	PIG-healthy1	30.56	15.18	12.6	85.2
	PIG-healthy2	34.42	15.29	10.6	86.2
	PIG-LVH1	40.01	26.51	14.3	165.2
	PIG-LVH2	55.23	35.34	13.9	178.2

Table 4.2: **Myocardial stiffness ($C_{\text{stiffness}}$) estimation across 2 clinical image-tracking cases and 8 FEA-simulated validation cases.** MC Dropout with 50 stochastic forward passes is used to compute a 95% CI. RE denotes relative error relative to reference values.

Source	Case	C (kPa) PINN	C Ref (kPa)	RE(%)↓	95% CI
Image tracking (Real)	Adult-1	0.095	0.10 [70]	5.0	[0.091, 0.100]
	Fetal-1	0.094	0.10 [97]	6.0	[0.089, 0.095]
FE tracking (Validation)	Adult-2	0.098	0.10	2.0	[0.094, 0.103]
	Adult-3	0.095	0.10	5.0	[0.091, 0.099]
	Fetal-2	0.098	0.10	2.0	[0.093, 0.103]
	Fetal-3	0.101	0.10	1.0	[0.096, 0.106]
	PIG-healthy1	0.089	0.09	1.1	[0.085, 0.093]
	PIG-healthy2	0.093	0.09	3.3	[0.089, 0.097]
	PIG-LVH1	0.377	0.40	5.8	[0.358, 0.396]
	PIG-LVH2	0.409	0.40	2.3	[0.389, 0.428]

Table 4.1 listed the physiological loading conditions (EDV, ESV, EDP and PSP) for the 2 image-tracking and 8 FEA-simulated validation cases. Tables 4.2 and 4.3 summarised the corresponding estimates of $C_{\text{stiffness}}$ and T_{max} . For the FEA cases, the ground truth reference values for $C_{\text{stiffness}}$ and T_{max} were those used in the FEA simulation settings, which were based on [70, 126, 128, 133, 134, 136]. For the image tracking cases, no ground truth was available, but the reference values were taken from the values or ranges of values reported in the literature [70, 97, 133–136]. For $C_{\text{stiffness}}$, we used values from previous FEA simulations that successfully replicated in vivo cardiac function and that used the same stiffness model as us. For adult T_{max} , a range of literature was used, including those performing inference from medical images and simulations that successfully replicated the cardiac function. For fetal T_{max} , literature on experimental measurements was used.

Table 4.3: **Active tension ($T_{\max}$) estimation across 2 clinical image-tracking cases and 8 FE-simulated validation cases.** MC Dropout with 50 stochastic forward passes provides a 95% CI. RE denotes relative error relative to reference values.

Source	Case	$T_{\max}$ (kPa) PINN	$T_{\max}$ Ref (kPa)	RE(%)↓	95% CI
Image tracking (Real)	Adult-1	124.8	95–200.7 [70, 133–135]	within	[112.6, 138.3]
	Fetal-1	36.9	23.9–59.2 [136]	within	[36.8, 41.7]
FE tracking (Validation)	Adult-2	103.0	100.3	2.73	[97.2, 108.1]
	Adult-3	100.3	100.3	0.05	[96.1, 104.5]
	Fetal-2	28.8	30.0	4.11	[27.1, 30.6]
	Fetal-3	38.8	40.0	3.02	[36.1, 41.6]
	PIG-healthy1	42.7	40.5	5.41	[39.5, 45.9]
	PIG-healthy2	38.6	40.5	4.62	[36.0, 41.2]
	PIG-LVH1	60.2	60.0	0.40	[57.1, 63.5]
	PIG-LVH2	58.6	60.0	2.26	[55.9, 61.3]

For the image tracking cases, the estimated $C_{\text{stiffness}}$ and $T_{\max}$ demonstrated reasonable agreement with these literature values, or fell within the literature range. For the FEA tracking cases, the estimated $C_{\text{stiffness}}$ and $T_{\max}$ showed strong agreement with the reference values, exhibiting relatively low RE across different subject groups. The RE for both $C_{\text{stiffness}}$ and $T_{\max}$ remained below 6.0% in all adult, fetal, and pig FEA cases, indicating reasonable predictive performance across a range of physiological conditions.

In addition, to quantify predictive uncertainty, we applied MC dropout at test time by enabling a dropout layer with a rate of $p = 0.1$ before the final regression head, and performed 50 stochastic forward passes per case to obtain a 95% CI for each predicted $C_{\text{stiffness}}$ and $T_{\max}$ (Tables 4.2 and 4.3). In all cases, the deterministic PINN prediction (with dropout disabled) lie within the corresponding 95% CI and typically close to the centre of the interval. Moreover, the CI widths, i.e. the differences between the upper and lower bounds, were small compared with the magnitude of the estimates themselves, indicating that the stochastic perturbations introduced by dropout led to only minor variations in $C_{\text{stiffness}}$ and $T_{\max}$. Together with the low relative errors, these narrow CIs supported the robustness and reliability of the proposed parameter estimators.

Compared to the traditional approach of performing iterative FEA simulations to back-compute $C_{\text{stiffness}}$ and $T_{\max}$, our PINN estimators achieved a tremendous reduction in computational time. A typical iterative FEA simulation required several hours for a single iteration (and one case usually required multiple iterations for converging), whereas our network completed the estimation in 63 ± 22 seconds per case.

Overall, the results demonstrated the reasonable accuracy of our $C_{\text{stiffness}}$ and $T_{\max}$ estimators, along with a significant reduction in computational cost compared to the traditional iterative FEA approach, highlighting the potential of *IMC-PINN-FE* as a real-time and clinically adoptable

solution for biomechanical parameter estimation.

4.3.2 Performance of PINN-FEA solver

4.3.2.1 Results of P-V loop

4

In this section, we compared the P-V loops obtained from whole cardiac cycle simulations, using the PINN-FEA solver (III of *IMC-PINN-FE*), the traditional volume-constrained FEA model, a state-of-the-art past LV PINN-FEA implementation by Buoso et al. [22], and a fixed image tracking approach. 2 representative image cases were used for this (Adult-2 and Fetal-2).

IMC-PINN-FE and volume-constrained FEA simulations were conducted as described in the methods section. The fixed image tracking method was essentially *IMC-PINN-FE* but where input displacement fields were fixed as obtained from the images, and were not subject to optimization. A neural network was trained solely to fit pressure values at each time point, using the fixed displacements (represented by motion mode amplitudes) as input. The loss function followed the same formulation as *IMC-PINN-FE* (Eq. 4.8), except that the volume constraint term ($\text{loss}_{\text{volume}}$) was excluded, since volume was already determined by the fixed displacements. This design isolated the pressure fitting process, allowing us to assess whether the unmodified image-derived motion fields, assumed to be error free, could satisfy the physical energy balance, and evaluated whether imaged displacements were consistent with biophysics.

Buoso et al.’s PINN-FEA adopted a PINN framework that predicted myocardial motion from pressure and active tension inputs. Unlike our method, which uses patient-specific motion fields to encode shape modes, Buoso et al. used multi-subject dataset and inter-patient shape variability to encode shape modes, and we used their published modes for implementing their approach. For inference, Buoso et al. used pressure and active tension waveforms derived from a Windkessel model to generate myocardial deformation over the cardiac cycle. At present, we did not implement a Windkessel model for *IMC-PINN-FE*, but relied on imaged LV volumes rather than Windkessel computations as inputs. To enable comparisons between the two methods, we amended Buoso et al.’s code to exclude the Windkessel model, using the FEA simulation pressures instead of pressure from the Windkessel model as inputs.

The results were shown in Fig. 4.5. In both image cases, *IMC-PINN-FE* achieved P-V loops (blue) that were nearly identical to those obtained from the traditional volume-constrained FEA simulations (orange), suggesting that *IMC-PINN-FE* effectively integrated biomechanics principles

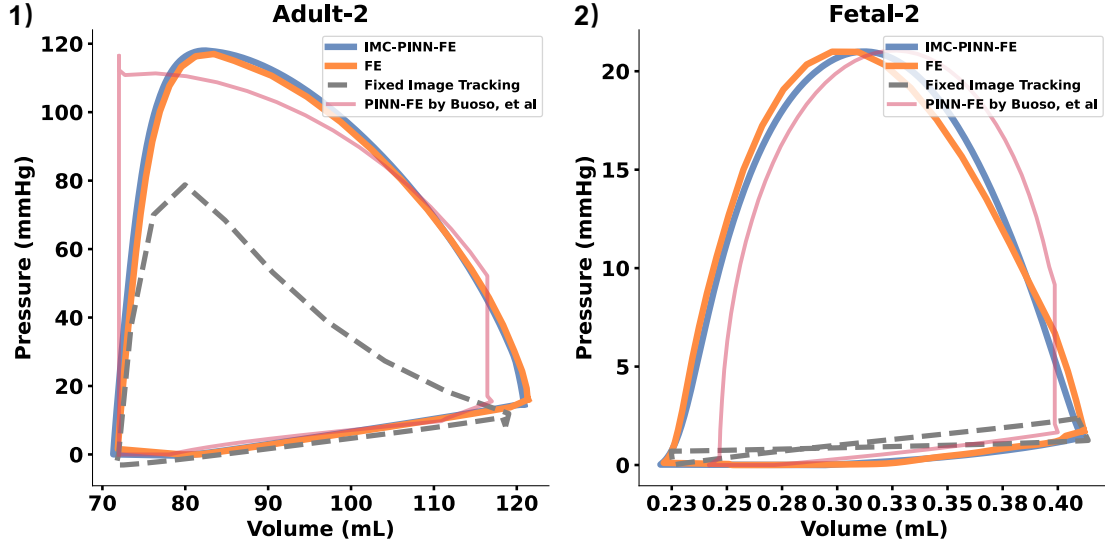


Figure 4.5: **P-V loop evaluation for (1) adult-2 and (2) fetal-2 cases**, comparing *IMC-PINN-FEA*, FEA simulation, fixed image tracking, and Buoso et al.’s *PINN-FEA* [22]. Blue represents the P-V loop predicted by *IMC-PINN-FE*, orange represents the P-V loop from FEA simulation, gray represents the P-V loop generated by fixed displacements from image tracking, and red represents the P-V loop predicted by Buoso et al.’s *PINN-FEA* [22].

despite maintaining effective data-driven constraints. Meanwhile, *IMC-PINN-FE* also enabled better match with imaged motion fields, as was demonstrated in Section 4.3.2.2.

In contrast, the fixed image tracking approach (gray) showed clear deviations, particularly in the systolic phase, where it significantly underestimated pressure compared to traditional FEA. This discrepancy arose because the network, constrained by fixed displacements, could not adjust myocardial motion during training to minimize both the PS pressure control loss (Eq. 4.6) and the energy balance loss (Eq. 4.4), leading to poor pressure estimates. This highlighted the importance of allowing the optimization of the imaged motions instead of directly adopting them, and that errors in image-derived displacements needed to be mitigated, as they could lead to inconsistency with biophysics laws.

The loss component comparisons among *IMC-PINN-FE*, traditional volume-constrained FEA, and fixed image tracking were presented in appendix 8.5.5. The results demonstrated that *IMC-PINN-FE* consistently achieved low physics losses, comparable to the FEA baseline, and were substantially lower than those of fixed tracking, which had considerable physical inconsistencies in pressure regulation and energy accumulation.

The *PINN-FEA* model proposed by Buoso et al. [22] also exhibited discrepancies compared to the traditional FEA reference P-V loop, particularly in the systolic volumes, which often deviated

from the FEA references. These deviations arose from the absence of explicit volume constraints to match the imaged volumes, and from the use of an algorithm to construct the isovolumetric phases despite deviations from the measured volumes (i.e., a pressure adjustment scheme that iteratively updated the pressure during the isovolumetric phases toward the target volumes). While this method facilitated the generation of smooth and plausible curves, it introduced artificial vertical truncations in the P-V loop and reduced adaptability to subject-specific variations. In contrast, *IMC-PINN-FE* simultaneously optimised both pressure through physics-based constraints, and displacements under volume constraints, enabling a more faithful reproduction of subject-specific P-V dynamics and better alignment with the imaged myocardial motion. However, since smoothing during image motion tracking removed constant volume features from the volume versus time curve, *IMC-PINN-FE* had the limitation of not explicitly demonstrating the isovolumetric phases.

4.3.2.2 Results of Geometric Deformation Accuracy

Fig. 4.6 presented box plots of DSC evaluation over 20 uniformly sampled time frames across the cardiac cycle for each of the 6 cases. DSC was used to quantify the spatial overlap between the predicted 3D LV shapes (from PINN or FEA simulations) and the image-derived ground truth (obtained via segmentation and *EFDL* motion tracking).

Across all cases, *IMC-PINN-FE* (blue) consistently achieved the highest DSC values (0.927 ± 0.039 across all 6 cases), demonstrating superior shape reconstruction accuracy and robustness. Traditional FEA simulations (orange) showed moderate performance, with wider spreads and lower medians, indicating greater deviation from image-derived shapes. Buoso et al.'s PINN-FEA [22] (red) yielded the lowest DSC values overall, with large variability and systematic underperformance across both adult and fetal cases. This reflected its limited capability for patient-specific shape recovery, due to reliance on generic shape modes and the absence of image-based constraints.

These findings were further illustrated in Fig. 4.7, which presented pointwise deformation error maps at 5 representative time points throughout the cardiac cycle for the Adult-3 case. The color encoded the magnitude of displacement error relative to the image-derived tracking, with red indicating large errors and blue indicating small errors. Traditional FEA showed substantial errors, particularly near the apex, where longitudinal motion was overestimated. The method by Buoso et al. exhibited even greater discrepancies from the ground truth, with widespread spatial errors across the myocardium. In contrast, *IMC-PINN-FE* consistently achieved the lowest deformation errors across the cycle, closely matching the image-derived motion.

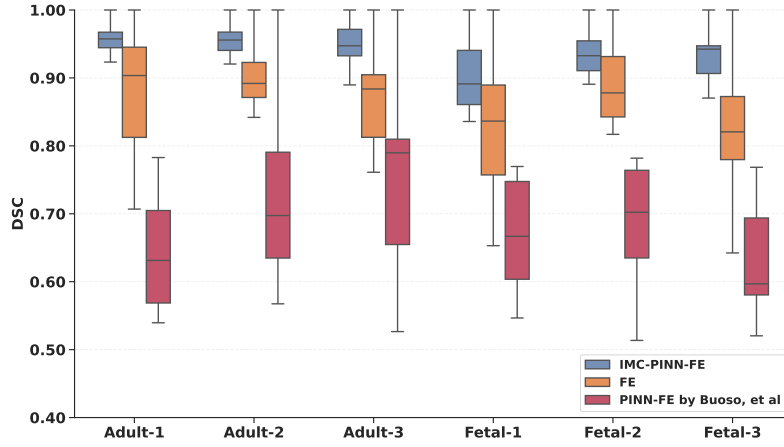


Figure 4.6: **Box plots of DSC evaluation** (20 uniformly sampled frames across the cardiac cycle for each of the 6 cases), comparing LV geometries from *IMC-PINN-FE*, traditional FEA, and Buoso et al.’s PINN-FEA [22] against image-derived results. Orange indicates the DSC between FEA outputs and image tracking, blue indicates *IMC-PINN-FE* vs. image tracking, and red indicates Buoso et al.’s PINN-FEA vs. image tracking.

This discrepancy stemmed from the lack of constraints to enforce a match of simulated cardiac motions to motions observed in the medical images. The traditional FEA modelling generated deterministic solutions based on predefined material properties and boundary conditions, and to enable accurate replication of patient-specific dynamics typically required extensive manual tuning of such input parameters. As for the PINN method by Buoso et al. [22], on top of not enforcing a match with imaged motions, its motion modes were not obtained from individual-specific dynamics, which may have contributed further to mismatch with imaged motions. In contrast, *IMC-PINN-FE* integrated physics-informed constraints with image-derived priors, enabling the model to adapt to personalised motion patterns and produce deformation fields that were both data-consistent and physiologically realistic.

4.3.2.3 Comparison of SOTA Models

Table 4.4: **Comparison of SOTA surrogate models evaluated on the adult-only test set.**

Model	Geometry	PV Error Metrics			Efficiency	
	DSC $\uparrow$	PSP (%) $\downarrow$	EDP (%) $\downarrow$	$R^2\uparrow$	Inf. (min) $\downarrow$	Param (M) $\downarrow$
MeshGraphNet	0.802 $\pm$ 0.020	2.8 $\pm$ 2.5	3.1 $\pm$ 2.0	0.98 $\pm$ 0.01	0.042 $\pm$ 0.005	2.292
GraphSAGE	0.831 $\pm$ 0.019	3.0 $\pm$ 2.1	2.7 $\pm$ 1.4	0.98 $\pm$ 0.01	0.034$\pm$0.001	0.810
GraphUNet	0.756 $\pm$ 0.033	10.2 $\pm$ 5.3	3.7 $\pm$ 2.4	0.95 $\pm$ 0.01	0.067 $\pm$ 0.010	3.256
FEA	0.883 $\pm$ 0.016	—	—	—	444.0 $\pm$ 15.0	—
Buoso’s PINN	0.701 $\pm$ 0.025	—	—	0.90 $\pm$ 0.02	2.333 $\pm$ 0.383	0.001
<i>IMC-PINN-FE</i> (ours)	0.952$\pm$0.011	0.6$\pm$0.3	1.0$\pm$0.8	0.99$\pm$0.01	5.833 $\pm$ 3.217	0.053

To the best of our knowledge, *IMC-PINN-FE* was the first framework that established a complete end-to-end pipeline connecting three-dimensional cardiac imaging, myocardial material pa-

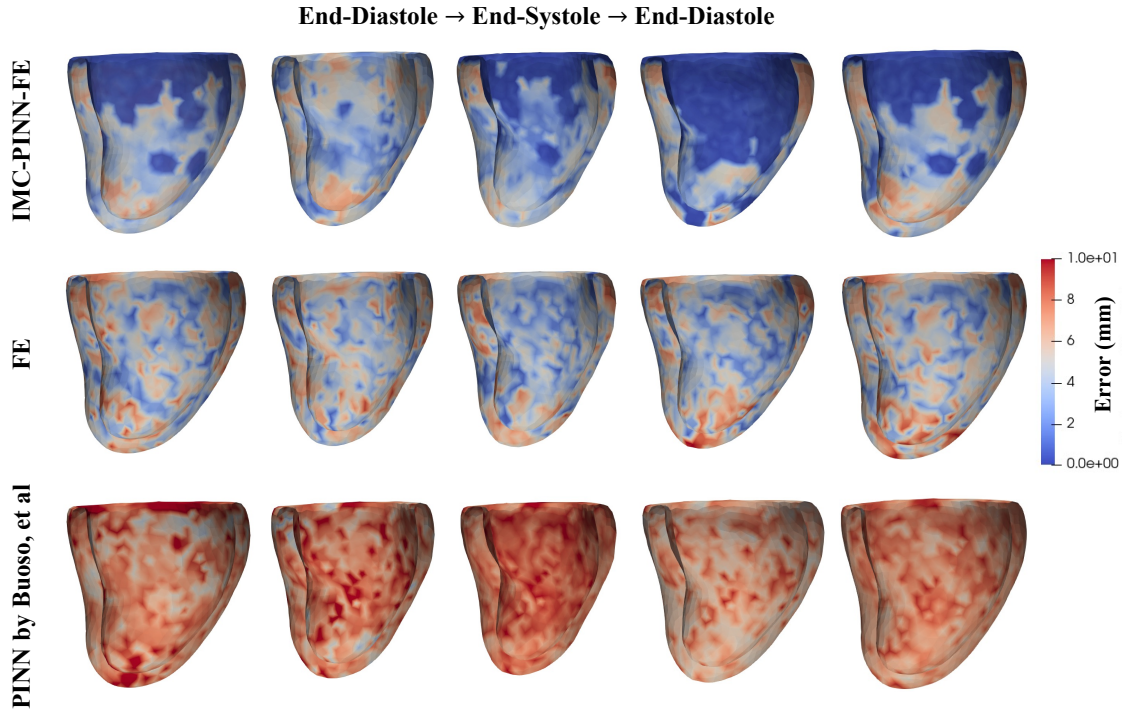


Figure 4.7: **Pointwise deformation error maps across different time points throughout the cardiac cycle (Adult-3).** Each row corresponds to a different method: *IMC-PINN-FE*, traditional FEA, and Buoso et al.’s PINN-FEA [22]. Color encodes the magnitude of pointwise error with respect to the image-derived motion (contour).

parameter estimation, and physically constrained full-cycle P-V prediction. Existing pre-trained DL surrogates, by contrast, could not adjust material parameters, did not model both diastolic and systolic phases, and were generally limited to diastolic-only deformation prediction [18–20].

To enable a rigorous and controlled comparison, we developed full-cycle implementations of 3 representative mesh-free GNN architectures, namely MeshGraphNet [137], GraphSAGE [138] and GraphUNet [139] (Table 4.4). Because no prior work provided publicly available models capable of representing an entire cardiac cycle, these architectures were trained from the ground up using the pre-computed FEA dataset described in appendix 8.6 to establish purely data-driven baselines for comparison. All evaluations were performed with fixed adult myocardial parameters and were restricted to an adult-only test set to ensure consistency. Traditional FEA simulation served as the gold standard for PV-related physiological accuracy, and *EFDL* tracking provided the reference standard for geometric accuracy.

In terms of DISP-DSC, *IMC-PINN-FE* (0.952 ± 0.011) substantially outperformed the purely data-driven surrogates (0.756–0.831) as well as the PINN-FEA model proposed by Buoso et al. (0.701). This improvement arose from the POD-mode constraint derived from real-image tracking,

which provided strong geometric regularisation. In comparison, FEA-only ground-truth models did not incorporate image-domain information and inevitably inherited modelling and numerical biases, making them more prone to local drift when generalised across meshes with different resolutions or topology.

With respect to P-V consistency, although time-dependent MeshGraphNet and GraphSAGE captured overall trends, their performance remained systematically inferior to *IMC-PINN-FE* across key physiological metrics, including PSP-RE, EDP-RE and PV R^2 . This indicated that purely data-driven surrogates generalised less effectively to unseen anatomical geometries compared with the patient-specific optimisation achieved by our PINN formulation.

In terms of computational efficiency, *IMC-PINN-FE* completed a full diastolic-to-systolic FEA-level surrogate prediction in approximately 5.8 minutes, which was about 75 times faster than traditional FEA simulation (444 minutes). Although slower than purely data-driven GNN surrogates (0.03–0.06 minutes), the proposed framework offered a more favourable balance between physical fidelity and computational speed while maintaining cycle-wise mechanical consistency.

Overall, the results demonstrated that *IMC-PINN-FE* provided superior geometric reconstruction, stronger physiological consistency, and more robust generalisation than existing deep-learning surrogates and PINN-FEA baselines.

4.4 Discussion

In this study, we introduced *IMC-PINN-FE*, a novel PINN framework that integrated image motion consistency with biomechanical modelling for patient-specific LV FEA simulations. Our method effectively addressed the limitations of traditional FEA models and existing PINN-based approaches, which often failed to sufficiently adhere to imaged motion and lacked physics-based estimation of myocardial stiffness and contractile function. The proposed framework was applicable to both MRI and echocardiography data. In essence, it enhanced the fidelity to imaged motion by incorporating physics constraints to derive more biophysically accurate myocardial dynamics.

In our experiments, we further observed that using shape modes for motion encoding improved convergence, whereas unencoded node-wise PINN-FEA methods such as [21] often suffered from convergence issues when applied to patient-specific rather than idealised geometries. The use of shape modes also significantly enhanced computational efficiency and reduced training time. For this reason, we adopted a reduced-order, motion-encoded PINN approach.

Compared to other reduced-order PINN frameworks, *IMC-PINN-FE* encoded motion based on patient-specific cardiac dynamics, rather than relying on inter-subject geometric variability from large multi-subject datasets. This strategy improved patient specificity and eliminated the need for extensive training datasets with ground-truth reconstructions, which were often difficult to obtain in clinical settings.

4

Consequent to the reduced-order modelling, our approach offered substantial computational advantages, reducing simulation time from over 7 hours with traditional iterative FEA models to under 10 minutes with *IMC-PINN-FE*, thereby improving its feasibility for clinical adoption. This efficiency also made it feasible for *IMC-PINN-FE* to support high throughput applications, such as building large simulation databases for deep learning training, or for large sample clinical testing.

In terms of performance, our framework enabled robust and efficient training by integrating image-informed constraints with physics-based governing equations. The inference results demonstrated high geometric accuracy (DSC: 0.927 ± 0.039) for image-derived LV shapes, while maintaining physiological consistency in the predicted P-V loops. Furthermore, the biomechanical parameter estimation component yielded reliable personalised predictions of myocardial stiffness and active tension, with relative errors of $2.81 \pm 1.64\%$ for $C_{\text{stiffness}}$ and $2.82 \pm 1.79\%$ for T_{max} .

Although *IMC-PINN-FE* used the same equation systems as traditional FEA modelling, it yielded different results with a better match to imaged motion, and yet produced a very similar P-V loop. This apparent discrepancy could be explained by the range of solutions that were possible under the FEA governing equations and by the difficulty in establishing the uniqueness of the solution given the large number of degrees of freedom. Essentially, *IMC-PINN-FE* was able to find another displacement solution that was close to the imaged motions but which still conformed to the physics equations. The key to enabling this was to not directly constrain *IMC-PINN-FE* to the imaged displacement fields, in line with the understanding that motion extraction from images had errors, but to allow it to have some deviations and space to find the closest solution conforming to both motion and physics constraints. Overall, *IMC-PINN-FE* offers an accurate and data-efficient alternative to conventional FEA methods, which were typically time-consuming.

4.4.1 Sensitivity and Error Analysis

In addition to the main quantitative evaluation, we also performed a set of complementary sensitivity analyses, with full details reported in appendix 8.5, to assess the robustness of *IMC-PINN-FE* to modelling and data uncertainties. First, we analysed the effect of the number of temporal frames

per cardiac cycle using the FEA-based validation case. *IMC-PINN-FE* remained stable over a wide range of frame counts, but very sparse sampling with fewer than approximately 8 frames per cycle led to noticeable deviations in the estimated stiffness, maximal tension and P-V loop. Around 20 frames per cycle provided a good compromise between phase resolution and computational cost, and was therefore adopted for both adult and fetal cases.

Second, we examined how the choice of unloaded reference affected the results by comparing the one-third diastolic approach to the inverse FEA unloading approach (appendix 8.5.2). The two references produced very similar material parameters and P-V loops in all groups, with somewhat larger but still moderate differences in the LVH pig models, indicating that the one-third diastolic is substantially similar to inverse FEA. A complementary volume perturbation study along the same diastolic filling trajectory showed that only larger deviations (over 10%) produced marked changes in EDP while PSP and the overall loop pattern remained relatively stable (appendix 8.5.3).

Third, we investigated sensitivity to imaging quality by adding synthetic Gaussian noise to the medical images before contour extraction and repeating the entire tracking and modelling pipeline (appendix 8.5.4). The dominant POD motion modes, the back-computed stiffness and maximal tension, and the reconstructed P-V loops were all robust for realistic noise levels and only started to degrade under deliberately severe perturbations.

In summary, these sensitivity studies supported that *IMC-PINN-FE* is reasonably robust to typical uncertainties in temporal sampling, unloaded-state definition and image-derived motion, while also delineating parameter ranges in which diastolic pressure estimates become more sensitive and should be interpreted with greater caution.

4.4.2 Limitations and Future Work

A practical limitation of the proposed framework is its dependency on the *EFDL* tracking, whose Fourier-based temporal representation and regularisation introduce an intrinsic smoothing of the motion. Very rapid transients, such as short isovolumic phases or highly irregular beats, may therefore be under-resolved, especially when the sampling is only approximately uniform over the cardiac cycle. Importantly, this under-resolution does not reflect a fundamental limitation of the *IMC-PINN-FE* coupling itself. As shown by the additional FEA experiment in appendix 8.5.6, when the same LV mechanics are driven by an idealised, Windkessel-constrained volume-time (V-t) trajectory with short near-constant-volume segments, *IMC-PINN-FE* faithfully reproduces the corresponding P-V loop and clearly preserves the isovolumetric phases. This indicates that the

blurring of isovolumetric plateaus in the main experiments arises primarily from the smoothed, image-derived V - t input rather than from the PINN-FEA formulation. In future work, this limitation can be mitigated by exploring alternative temporal parameterisations, such as non-uniform or adaptive time bases that allocate higher effective resolution to fast transients while using coarser sampling in slowly varying phases, or by replacing or complementing *EFDL* with more efficient, modality-specific tracking or segmentation networks (such as Co-attentionSTN [100]) when larger, site-specific datasets are available. At the present stage, however, the choice of *EFDL* is a deliberate trade-off tailored to data-scarce, patient-specific research scenarios. It provides cross-modality applicability (MRI and echocardiography) without requiring large annotated motion datasets or extensive retraining for each new centre or scanner.

A further limitation is that the present validation still does not include direct invasive or ex-vivo reference measurements of myocardial properties. For the clinical image-tracking cases, only non-invasive imaging and literature-based physiological ranges are available, and the FEA validation cases are constrained by the assumed material laws used to generate them. In future work, we plan to extend the framework to cohorts where invasive catheter pressure recordings and, for animal models such as pigs, ex-vivo biomechanical testing are available, so that the estimated stiffness and active tension can be compared more directly against independently measured tissue properties.

The hyperparameter settings were selected based on preliminary hyperparameter optimization and prior work on PINNs, aiming to balance expressive power and training stability, and the same hyperparameters were used across all cases without subject-specific tuning. As this preliminary tuning was stopped once reasonable performance was achieved, we acknowledge that further hyperparameter optimization may still yield additional improvements. In future work, we will adapt these settings to different imaging modalities, noise levels, and physiological regimes to better match scenario-specific requirements.

4.5 Conclusion

In this work, we presented *IMC-PINN-FE*, a patient-specific PINN framework for LV biomechanical modelling that integrates image-derived motion with biomechanical constraints from FEA. Its main strength is that it provides a practical and physics-consistent approach for recovering patient-specific myocardial stiffness and peak active tension from routine cine MRI and Echo, without relying on costly large-scale supervised training. By encoding motion using POD modes

and constraining the solution through biomechanical consistency, the framework generated LV P-V loops that closely matched conventional FEA simulations while also improving consistency with observed image motion. Across adult, fetal, and porcine studies, *IMC-PINN-FE* reduced computation from several hours to under ten minutes per case and maintained low REs in the recovered parameters. The sensitivity studies additionally showed that the method remained reasonably stable under practical perturbations in temporal sampling, unloaded-state definition, and image noise.

However, these advantages come with clear limitations. The framework still depends on patient-specific optimisation for each new case, which limits throughput and prevents near-real-time deployment. Its POD-based representation improves efficiency, but also restricts the deformation field to a reduced-order space, which may not capture richer spatially varying mechanics in more complex anatomies. In addition, the framework is designed mainly for personalised inverse estimation rather than for building a reusable surrogate across many subjects, and therefore does not fully address the need for geometry-generalised and cycle-wide prediction. For this reason, although *IMC-PINN-FE* provides an important step toward clinically usable image-constrained biomechanics, it also highlights the need for a different class of models that can learn directly from mesh structure, generalise across anatomies, and provide much faster inference once trained.

These considerations motivated the next stage of this thesis, namely the development of graph-based biomechanical surrogates. Compared with the PINN-based formulation presented here, graph networks offer a more natural way to represent irregular cardiac meshes, avoid explicit restriction to a reduced POD motion space, and enable pre-trained models that can be reused across subjects.

Graph Neural Networks for Left Ventricular Unloading

S. Mu et al., “Heartunloadnet: A cycle-consistent graph network with reduced supervision for predicting unloaded cardiac geometry from diastolic states,” *Computer Methods and Programs in Biomedicine*, p. 109 243, 2026

Contents

5.1 Introduction	73
5.1.1 Contributions	75
5.2 Methods	75
5.2.1 Finite Element Ground Truth Generation	75
5.2.2 HeartUnloadNet Architecture	76
5.2.3 Experimental Setup	81
5.3 Experiments and Results	85
5.3.1 Quantitative Performance Comparison	85
5.3.2 Reduced Supervised Learning Performance	86
5.3.3 Ablation Studies	87
5.3.4 Generalization to Unseen Global Physiological Values	89
5.3.5 Qualitative Visualization of Predictions	90
5.3.6 Advantages over PCA-Based and Inverse FEA Approaches	91
5.4 Discussion	92
5.4.1 Novelty and Contribution	92
5.4.2 Implications	93
5.4.3 Architectural Strengths and Outlook	93
5.4.4 Limitations and Future Work	94
5.5 Conclusion	96

The unloaded cardiac geometry (i.e., the state of the heart devoid of luminal pressure) serves as a valuable zero-stress and zero-strain reference and is critical for personalized biomechanical

modelling of cardiac function, to understand both healthy and diseased physiology and to predict the effects of cardiac interventions. However, estimating the unloaded geometry from clinical images remains a challenging task. Traditional inverse FEA approaches in cardiac biomechanics iteratively estimate unloaded geometries or biomechanical parameters from observed physiological states, rather than directly solving the forward mechanical problem, and are therefore computationally expensive due to the repeated optimisation process [78]. In this chapter, we introduce *HeartUnloadNet*, a DL framework that predicts the unloaded LV shape directly from the ED mesh while explicitly incorporating biophysical priors. The network accepts a mesh of arbitrary size along with physiological parameters such as ED pressure, myocardial stiffness scale, and fiber helix orientation, and outputs the corresponding unloaded mesh. It adopts a graph attention architecture and employs a cycle-consistency strategy to enable bidirectional (loading and unloading) prediction, allowing for partial self-supervision that improves accuracy and reduces the need for large training datasets. Trained and tested on 10,350 FEA simulations across diverse LV geometries and physiological conditions, *HeartUnloadNet* achieves sub-millimeter accuracy, with a DSC of 0.986 ± 0.023 and a HD of 0.083 ± 0.028 cm, while reducing inference time to just 0.02 seconds per case, over $10^5 \times$ faster and significantly more accurate than traditional inverse FE solvers. Ablation studies confirm the effectiveness of the architecture. Notably, the cycle-consistent design enables the model to maintain a DSC of 97% even with as few as 200 training samples. This work thus presents a scalable and accurate surrogate for inverse FEA solvers, supporting real-time clinical applications in the future.

5.1 Introduction

Myocardial biomechanics studies such as image-based FEA [2, 3] and myocardial stiffness estimation from medical images [23] are important tools that have improved our understanding of cardiac physiology, and have significant potential for future translation into clinical tools for disease evaluation and surgical planning.

In such studies, the unloaded configuration of the human LV, defined as a theoretical zero pressure, low stress state of the LV in the absence of intracardiac fluid pressure, serves as a fundamental reference. It underlies the computation of myocardial stress and strain, initializes FEA simulations, and supports the personalization of material properties across varying loading conditions [140, 141]. However, this state is clinically unobservable, as the heart is continuously subjected to intracardiac pressure [141]. Consequently, medical imaging can only capture pressure-

loaded geometries, such as those observed at ED [141]. Developing efficient algorithms to estimate the unloaded state from cardiac geometries observable in medical images is therefore highly valuable for biomechanical modelling.

Current strategies for reconstructing the unloaded cardiac geometry rely on performing inverse FEA using the ED cardiac geometry reconstructed from medical images [12, 142, 143]. This remains the gold standard due to its strong biophysical foundations. However, it is computationally expensive, typically requiring several hours to solve a single case with a patient-specific LV mesh containing more than 1,500 nodes [78]. Moreover, inverse FEA approaches exhibit limited accuracy, particularly when confronted with poor-quality meshes or anatomical irregularities [144].

To address these limitations, the development of supervised DL models for estimating the unloaded cardiac state could offer significant advantages by drastically reducing computational time. Prior studies have demonstrated the feasibility of estimating the unloaded state for non-cardiac vascular structures such as the aorta [145, 146], but to date, no such attempts have been reported for the heart. For clinical and research applications, it is essential that such DL models be capable of handling variations in cardiac mesh topology, myocardial helix angle configuration, tissue stiffness, and in vivo blood pressure. Furthermore, given the high cost of generating large-scale FE simulation ground truths, methods that can reduce the required dataset size while maintaining performance would be particularly beneficial.

In this chapter, we introduce a GNN framework for accurately estimating the unloaded state of input cardiac meshes across varying helix angles, myocardial stiffness values, and ED pressure conditions. Unlike standard MLP, which require fixed input mesh sizes and consistent node ordering, the GNN inherently encodes node connectivity and spatial relationships, allowing the model to generalize across hearts with diverse mesh resolutions and topologies [147]. This topology-agnostic design eliminates the need for pre-processing steps such as template registration or node indexing alignment, enabling direct analysis of meshes with heterogeneous structures.

To reduce the amount of FEA-generated ground-truth supervision required (i.e., the number of labelled unloading simulations), we adopt a cycle-consistent training strategy for reduced supervision, inspired by CycleGAN [148]. Specifically, we enforce that the predicted unloaded geometry must accurately reconstruct the original ED shape through a reverse deformation path. This physiological reversibility provides an implicit supervisory signal, enabling the network to learn meaningful deformation mappings in a reduced supervised setting.

Furthermore, to fully replicate the biophysical context of inverse FEA modelling, we incorporate

global material parameters, such as fiber orientation and myocardial stiffness, into the network. These parameters are provided as global conditioning inputs, allowing the model to adapt its deformation predictions to patient-specific physiological conditions. By embedding such biophysical priors, our framework preserves the interpretability and material sensitivity of traditional inverse FEA approaches, while retaining the computational efficiency and scalability of DL. These design choices collectively define *HeartUnloadNet*, our proposed method for predicting unloaded cardiac geometry.

5.1.1 Contributions

Our main contributions are fourfold:

1. We develop a graph attention-based network to compute the load free state of the heart from an image-derived loaded state, which achieves state-of-the-art accuracy, with a mean DSC exceeding 98% and an average HD below 0.85 mm, outperforming previously reported DL-based and inverse FEA-based methods to the best of our knowledge.
2. Our graph attention-based autoencoder naturally accommodates diverse mesh topologies across patients, enabling robust deformation modeling without requiring fixed templates or mesh alignment pre-processing.
3. We introduce a cycle-consistent learning mechanism that enables effective training under a reduced supervised regime. Notably, using only around 200 labelled cases (approximately 3% of all samples), our model achieves 97% DSC accuracy on unseen test data, highlighting its data efficiency and generalization capability.
4. Our network is conditioned on global biophysical parameters, including fiber orientation and myocardial stiffness, and on in vivo pressures, and as such unloaded state predictions can be performed for varied combinations of these parameters.

5.2 Methods

5.2.1 Finite Element Ground Truth Generation

Since the *HeartUnloadNet* framework is data-driven, we first establish paired ground truth unloaded and ED LV myocardial mesh geometries, which we obtain via a traditional FEA model described

in chapter 2.3. Only the passive stiffness component is considered for our task of diastolic loading, the active contractile component is ignored.

In the process of obtaining the ground truth, we fix the material parameters b_{ij} . The only global parameters allowed to vary are the LV pressure P_{cavity} , the global stiffness parameter C , and the two helix angles θ_{endo} and θ_{epi} . By prescribing these four variables, the FEA model minimizes the potential energy using a Newton–Raphson solver starting from the unloaded state, the Newton iterations converge to the equilibrium ED configuration under the applied LV pressure P_{cavity} .

5

5.2.2 HeartUnloadNet Architecture

5.2.2.1 Overview

HeartUnloadNet is a cycle-consistent graph attention-based framework designed to infer the unloaded LV geometry from an observed ED state. As illustrated in Fig. 5.1 and Alg. 1, the architecture integrates both local geometric structures and global physiological parameters through two main components:

- A **graph fusion encoder** (Graph Attention Encoder and Global Parameter Encoder) that integrates node-level mesh features with global physiological parameters using a structured attention mechanism.
- A **cycle-consistent bi-directional decoder** that estimates nodal displacements in both directions, from the ED state to the unloaded configuration and back, using a shared encoder with two direction-specific decoders to enforce cycle consistency and structural coherence.

5.2.2.2 Graph Fusion Encoder

The encoder comprises three main components: Graph attention-based mesh encoder, global parameter encoder, and cross-attention fusion.

Graph attention-based encoder (Fig. 5.2): Given an input mesh with node features $\mathbf{H}^{(0)}$, the encoder applies three sequential graph attention layers. Each graph attention layer updates node features by aggregating neighbouring node information through attention weights [149]:

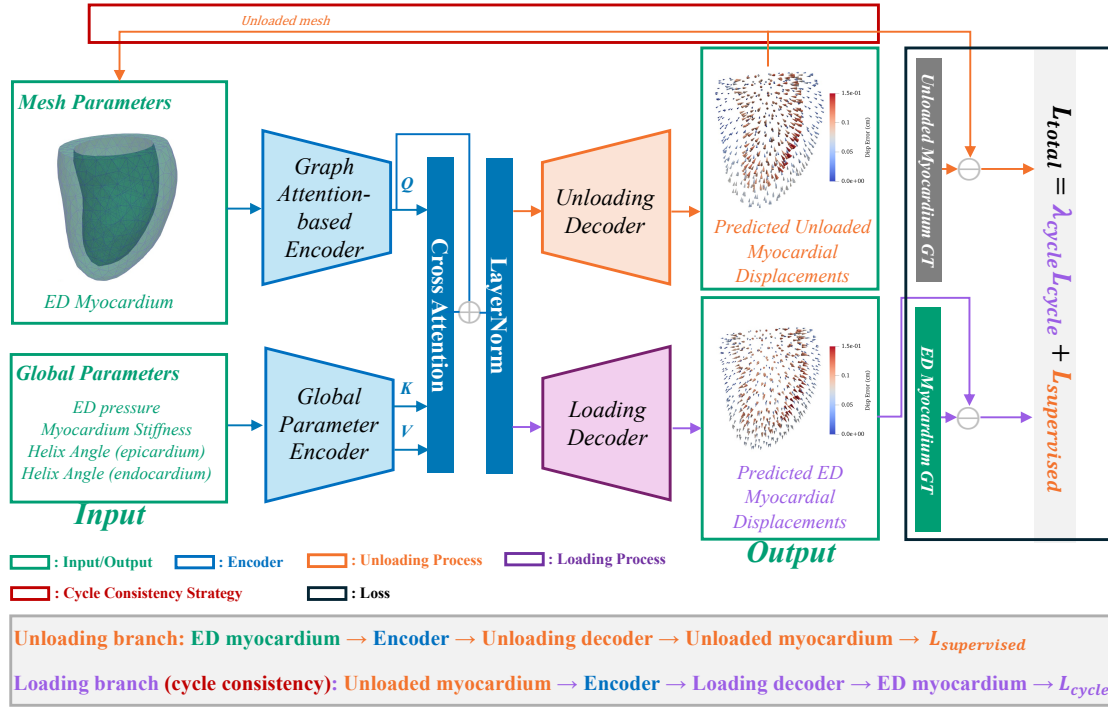


Figure 5.1: **Schematic of the proposed *HeartUnloadNet* for estimating the unloaded LV configuration from an ED state.** Inputs comprise the ED myocardial tetrahedral mesh (“Mesh parameters”) and global physiological parameters (“Global parameters”, including ED pressure, Fung stiffness scaling coefficient and helix angles). A shared graph attention-based encoder processes the mesh, while a global-parameter encoder embeds the scalar inputs, their features are fused by cross-attention and layer normalisation, yielding a node-wise representation conditioned on both local geometry and global physiology. On the unloading workflow, a decoder maps the fused features to nodal displacements and the corresponding unloaded myocardium, which are constrained by a supervised loss $L_{supervised}$ against the FEA inverse-unloading reference. On the loading workflow, the unloaded configuration is passed through the same encoder and a loading decoder to reconstruct the ED myocardium, with a cycle-consistency loss L_{cycle} enforcing agreement with the ED ground truth. The network is trained by minimising $L_{total} = \lambda_{cycle} L_{cycle} + L_{supervised}$. Colours indicate functional roles: green for inputs/outputs, blue for encoders and feature fusion, orange and purple for unloading and loading decoders, red for the cycle-consistency path, and black boxes for loss terms.

$$\begin{aligned} \mathbf{H}^{(\ell)} &= \text{GAT}^{(\ell)}(\mathbf{H}^{(\ell-1)}) + \mathbf{H}^{(\ell-1)} \\ &= \left\|_{m=1}^M \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(m, \ell)} \mathbf{W}^{(m, \ell)} \mathbf{H}_j^{(\ell-1)} \right) + \mathbf{H}^{(\ell-1)} \in \mathbb{R}^{n \times d}, \quad \ell = 1, 2, 3 \end{aligned} \quad (5.1)$$

$$\alpha_{ij}^{(m, \ell)} = \frac{\exp \left(\text{LeakyReLU} \left(\mathbf{a}^{(m, \ell) \top} \left[\mathbf{W}^{(m, \ell)} \mathbf{H}_i^{(\ell-1)} \parallel \mathbf{W}^{(m, \ell)} \mathbf{H}_j^{(\ell-1)} \right] \right) \right)}{\sum_{k \in \mathcal{N}(i)} \exp \left(\text{LeakyReLU} \left(\mathbf{a}^{(m, \ell) \top} \left[\mathbf{W}^{(m, \ell)} \mathbf{H}_i^{(\ell-1)} \parallel \mathbf{W}^{(m, \ell)} \mathbf{H}_k^{(\ell-1)} \right] \right) \right)} \quad (5.2)$$

where $\mathbf{H}^{(\ell)} \in \mathbb{R}^{n \times d}$ is the node feature matrix at the ℓ -th Graph Attention Network (GAT) layer, with each row $\mathbf{H}_i^{(\ell)}$ representing node i . Here, n denotes the number of mesh nodes, and d is

the dimensionality of node features. In our setting, $\mathbf{H}^{(0)}$ is initialized using the normalized 3D coordinates from the input ED mesh $\widetilde{\mathcal{M}}_{\text{ED}}$. The operator $\text{GAT}^{(\ell)}$ refers to a multi-head Graph Attention Network layer that aggregates information from neighbouring nodes via learned attention weights. The concatenation operator $\|_{m=1}^M$ denotes the output combination across M independent attention heads, each followed by a nonlinear activation function $\sigma(\cdot)$, instantiated as ReLU in our case. The set $\mathcal{N}(i)$ denotes the 1-hop neighbours of node i , implicitly encoding the mesh connectivity. The term $\alpha_{ij}^{(m,\ell)}$ represents the normalized attention coefficient associated with edge (i, j) in the m -th attention head at layer ℓ , reflecting the relative importance of node j to node i . These coefficients are computed using a softmax operation over all neighbours $j \in \mathcal{N}(i)$, as described in Eqn 5.2. The attention parameters $\mathbf{a}^{(m,\ell)}$ and transformation matrices $\mathbf{W}^{(m,\ell)}$ are learnable for each attention head and layer. The addition of $\mathbf{H}^{(\ell-1)}$ serves as a residual connection, enabling the network to preserve low-level features and improving training stability by mitigating vanishing gradients.

After the final GAT layer, node-level features $\mathbf{H}^{(3)}$ are aggregated into a global mesh-level representation $\mathbf{g}_{\text{mesh}}$ via mean pooling [150]. This global vector encodes the overall geometry and feature distribution of the entire mesh, providing essential context for subsequent decoding. By incorporating this global information, the decoder is expected to better capture whole-organ structural properties and implicitly enforce FE-like constraints, such as volume preservation and global energy consistency.

$$\mathbf{g}_{\text{mesh}} = \frac{1}{n} \sum_{i=1}^n \mathbf{H}_i^{(3)} \in \mathbb{R}^d \quad (5.3)$$

where n is the number of nodes in the mesh, and $\mathbf{g}_{\text{mesh}}$ is then broadcast back to each node to match the node-level dimension for subsequent fusion operations.

To incorporate global context into each node's local representation, we employ a cross-attention mechanism [151], where the mesh node features $\mathbf{H}^{(3)}$ serve as queries $\mathbf{Q}_m^{\mathbf{H}}$, and the broadcasted global vector $\mathbf{g}_{\text{mesh}}$ provides keys $\mathbf{K}_m^{\mathbf{g}_{\text{mesh}}}$ and values $\mathbf{V}_m^{\mathbf{g}_{\text{mesh}}}$. Specifically, the fused representation is computed as:

$$\mathbf{Z}_{\text{mesh}} = \text{LayerNorm} \left(\mathbf{H}^{(3)} + \left\|_{m=1}^M \text{softmax} \left(\frac{\mathbf{Q}_m^{\mathbf{H}} (\mathbf{K}_m^{\mathbf{g}_{\text{mesh}}})^\top}{\sqrt{d_m}} \right) \mathbf{V}_m^{\mathbf{g}_{\text{mesh}}} \right) \right) \in \mathbb{R}^{n \times d} \quad (5.4)$$

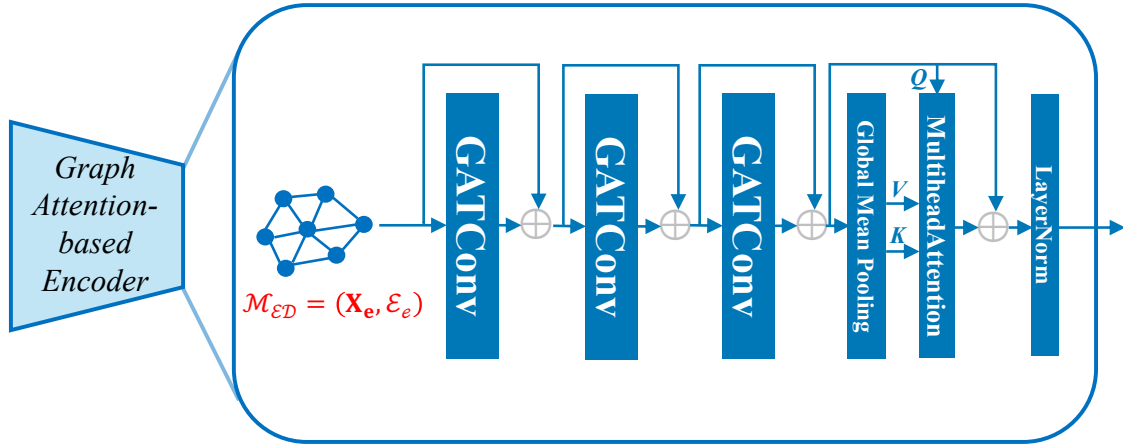


Figure 5.2: **Architecture of the graph attention-based encoder used in *HeartUnloadNet*.** The input is the ED myocardial mesh represented as a graph $\mathcal{M}_{\text{ED}} = (\mathbf{X}_e, \mathcal{E}_e)$, where $\mathbf{X}_e$ are node features and $\mathcal{E}_e$ the edge set. Node features are successively updated by three GATConv layers with residual connections (“ $\oplus$ ”). A global mean pooling operation then aggregates the node-wise features into a graph-level descriptor, which is combined with the node features in a MultiheadAttention block using query (Q), key (K) and value (V) projections. The output of this attention module, with an additional residual connection and layer normalisation, provides the fused latent representation of the ED myocardium used by the downstream decoders.

$$\begin{aligned}
 \mathbf{Q}_m^{\text{H}} &= \mathbf{H}^{(3)} \mathbf{W}_Q^{(m)} \in \mathbb{R}^{n \times d_m}, \\
 \mathbf{K}_m^{\text{g}_{\text{mesh}}} &= \mathbf{g}_{\text{mesh}} \mathbf{W}_K^{(m)} \in \mathbb{R}^{1 \times d_m}, \\
 \mathbf{V}_m^{\text{g}_{\text{mesh}}} &= \mathbf{g}_{\text{mesh}} \mathbf{W}_V^{(m)} \in \mathbb{R}^{1 \times d_m}
 \end{aligned} \tag{5.5}$$

where $\mathbf{H}^{(3)} \in \mathbb{R}^{n \times d}$ is the node feature matrix after the third GAT layer, where each row $\mathbf{H}_i^{(3)}$ corresponds to node i . The operator $\text{LayerNorm}(\cdot)$ denotes layer normalization applied across the feature dimension, which normalizes each node’s feature vector independently to stabilize training and improve convergence [152]. The matrices $\mathbf{W}_Q^{(m)}, \mathbf{W}_K^{(m)}, \mathbf{W}_V^{(m)} \in \mathbb{R}^{d \times d_m}$ are learnable projection matrices for the m -th attention head, where d_m is the feature dimensionality of each head.

This step allows each node to selectively attend to different aspects of the global mesh embedding, enabling feature modulation conditioned on the overall geometry. Such cross-attention enhances the network’s capacity to model dependencies beyond local neighbourhoods, crucial for biomechanical systems, where global shape and boundary-aware context jointly influence tissue behaviour.

Global Parameter Encoder: To incorporate global physiological priors into the model, we encode the pressure–stiffness–orientation tuple $(P, C, \theta_{\text{endo}}, \theta_{\text{epi}})$ into a compact feature vector via a MLP:

$$\mathbf{g}_{\text{global}} = \text{MLP}^{\text{global}}([P, C, \theta_{\text{endo}}, \theta_{\text{epi}}]) \in \mathbb{R}^d$$

It is worth noting that each of the four physiological parameters is normalized to the range $[0, 1]$ based on its respective physical bounds prior to encoding, with detailed values provided in the following section. This normalization promotes stable training and ensures balanced contributions from each parameter.

Cross-Attention Fusion: The mesh-level features $\mathbf{Z}_{\text{mesh}}$ and global physiological parameters $(P, C, \theta_{\text{endo}}, \theta_{\text{epi}})$ are integrated via a cross-attention fusion:

$$\mathbf{Z}_{\text{Fusion}} = \text{LayerNorm} \left(\mathbf{Z}_{\text{mesh}} + \parallel_{m=1}^M \text{softmax} \left(\frac{\mathbf{Q}_m^{\text{Z}_{\text{mesh}}} (\mathbf{K}_m^{\text{g}_{\text{global}}})^\top}{\sqrt{d_m}} \right) \mathbf{V}_m^{\text{g}_{\text{global}}} \right) \quad (5.6)$$

where $\mathbf{Q}_m^{\text{Z}_{\text{mesh}}}$, $\mathbf{K}_m^{\text{g}_{\text{global}}}$, and $\mathbf{V}_m^{\text{g}_{\text{global}}}$ are the query, key, and value matrices for the m -th attention head, computed from $\mathbf{Z}_{\text{mesh}}$ and $\mathbf{g}_{\text{global}}$, respectively. The cross-attention mechanism enables each node to selectively attend to the global physiological context.

5.2.2.3 Cycle-Consistent Bi-Directional Decoder

To achieve optimal performance while minimizing the required number of ground truth samples, we propose a cycle-consistent bi-directional decoder that estimates nodal displacements in both directions: ED to unloaded ($e \rightarrow u$) and unloaded to ED ($u \rightarrow e$). This is inspired by the principle of CycleGAN [148], where an image-to-image translation network is trained bidirectionally, with cycle consistency constraints imposed to ensure that an image can be faithfully reconstructed after a forward and subsequent backward translation. This cycle consistency provides a form of reduced supervision that regularizes learning, enhances generalization, and mitigates overfitting, thereby improving performance and reducing the required labelling sample size. Here, we introduce a similar reconstruction loop that enables self-supervised learning: the network predicts the unloaded mesh from the ED state, then reconstructs the ED mesh from the predicted unloaded state. Minimizing the discrepancy between the cyclic reconstructed and original ED mesh provides this additional self-supervision.

Formally, a shared encoder is employed for both directions, ED to unloaded and unloaded back to ED, to extract a unified fused feature representation $\mathbf{Z}_{\text{Fusion}}$. Two distinct MLP decoders are

subsequently used to predict displacements in each direction from $\mathbf{Z}_{\text{Fusion}}$:

$$\widehat{\Delta \mathbf{u}}_{e \rightarrow u} = \text{MLP}^{e \rightarrow u}(\text{Dropout}(\mathbf{Z}_{\text{Fusion}})) \in \mathbb{R}^{n \times 3}, \quad \widehat{\mathcal{M}}_u = \widetilde{\mathcal{M}}_{\text{ED}} + \widehat{\Delta \mathbf{u}}_{e \rightarrow u} \quad (5.7)$$

$$\widehat{\Delta \mathbf{u}}_{u \rightarrow e} = \text{MLP}^{u \rightarrow e}(\text{Dropout}(\mathbf{Z}_{\text{Fusion}})) \in \mathbb{R}^{n \times 3}, \quad \widehat{\mathcal{M}}_{\text{ED}} = \widehat{\mathcal{M}}_u + \widehat{\Delta \mathbf{u}}_{u \rightarrow e} \quad (5.8)$$

where $\widehat{\Delta \mathbf{u}}_{e \rightarrow u}, \widehat{\Delta \mathbf{u}}_{u \rightarrow e} \in \mathbb{R}^{n \times 3}$ are the predicted displacement vectors at each node. $\widetilde{\mathcal{M}}_{\text{ED}}$ denotes the input ED mesh geometry, and $\widehat{\mathcal{M}}_u, \widehat{\mathcal{M}}_{\text{ED}}$ are the predicted unloaded and reconstructed ED meshes, respectively. $\text{Dropout}(\cdot)$ applies dropout regularization to the fused features during training, helping to prevent overfitting by randomly deactivating a fraction of the feature dimensions [153].

5.2.2.4 Loss Function

Unlike the physics-constrained framework in Chapter 4, the model in Chapter 5 does not explicitly enforce volume preservation or global energy consistency. Instead, the network is trained using FEM-generated displacement and pressure supervision together with bidirectional cycle-consistency constraints to encourage temporally coherent biomechanical predictions. Based on the predicted meshes in both directions, the total loss is defined as:

$$\mathcal{L}_{\text{total}} = \left\| \widehat{\mathcal{M}}_u - \mathcal{M}_u \right\|^2 + \lambda_{\text{cycle}} \cdot \left\| \widehat{\mathcal{M}}_{\text{ED}} - \widetilde{\mathcal{M}}_{\text{ED}} \right\|^2$$

where λ_{cycle} controls the strength of the cycle-consistency constraint.

5.2.3 Experimental Setup

5.2.3.1 Data Acquisition

We constructed a dataset of LV geometries based on a statistical shape model derived from PCA of the *Cardiac Atlas Project (CAP)* (<https://www.cardiacatlas.org/left-ventricular-modes/>). The model was trained on 1,991 asymptomatic subjects from the MESA cohort [154] and captures dominant modes of anatomical variation at ED, including size, sphericity, and concentricity. To synthesize anatomically plausible ED geometries, we sampled PCA weights from a standard normal distribution and truncated each mode to $[-2, +2]$, corresponding to approximately 95% of population-level variation due to the zero mean and unit variance of each mode. We randomly

Algorithm 1: HeartUnloadNet

Input : Global physiological parameters $(P_{cavity}, C, \theta_{endo}, \theta_{epi})$, ED myocardial mesh $\widetilde{\mathcal{M}}_{ED} = (\mathbf{X}_e, \mathcal{E}_e)$

Output: Predicted displacement $\widehat{\Delta \mathbf{u}}_{e \rightarrow u}$

Ground Truth from FEA Simulation:

Given global physiological parameters $(P_{cavity}, C, \theta_{endo}, \theta_{epi})$, we apply forward FEA loading from $\mathcal{M}_u$ to obtain:

- Ground-truth ED mesh: $\widetilde{\mathcal{M}}_{ED} = \mathcal{M}_u + \Delta \mathbf{u}_{u \rightarrow e}$
- Supervised pair: $(\mathcal{M}_u, \widetilde{\mathcal{M}}_{ED})$

Training Process;

for $epoch = 1$ **to** T **do**

for $\mathbf{H}^{(0)} \in \{\widetilde{\mathcal{M}}_{ED}, \widetilde{\mathcal{M}}_u\}$ **do**

Encoder

 Graph attention-based encoder:

$$\mathbf{H}^{(\ell)} = \text{GAT}^{(\ell)}(\mathbf{H}^{(\ell-1)}) + \mathbf{H}^{(\ell-1)}, \quad \ell = 1, 2, 3$$

$$\mathbf{g}_{\text{mesh}} = \text{MeanPool}(\mathbf{H}^{(3)}) \in \mathbb{R}^d \quad (\text{broadcast to } \mathbb{R}^{n \times d})$$

$$\mathbf{Z}_{\text{mesh}} = \text{LayerNorm} \left(\mathbf{H}^{(3)} + \left\|_{m=1}^M \text{softmax} \left(\frac{\mathbf{Q}_m^H (\mathbf{K}_m^{\text{g}_{\text{mesh}}})^\top}{\sqrt{d_m}} \right) \mathbf{V}_m^{\text{g}_{\text{mesh}}} \right) \right) \in \mathbb{R}^{n \times d}$$

 Global parameter encoder: $\mathbf{g}_{\text{global}} = \text{MLP}^{\text{global}}([P, C, \theta_{endo}, \theta_{epi}])$

 Cross-Attention Fusion:

$$\mathbf{Z}_{\text{Fusion}} = \text{LayerNorm} \left(\mathbf{Z}_{\text{mesh}} + \left\|_{m=1}^M \text{softmax} \left(\frac{\mathbf{Q}_m^{\text{Z}_{\text{mesh}}} (\mathbf{K}_m^{\text{g}_{\text{global}}})^\top}{\sqrt{d_m}} \right) \mathbf{V}_m^{\text{g}_{\text{global}}} \right) \right) \in \mathbb{R}^{n \times d}$$

Decoder

 Predicted unloaded displacements (with $\mathbf{H}^{(0)} \leftarrow \widetilde{\mathcal{M}}_{ED}$):

$$\widehat{\Delta \mathbf{u}}_{e \rightarrow u} = \text{MLP}^{e \rightarrow u}(\text{Dropout}(\mathbf{Z}_{\text{Fusion}})) \in \mathbb{R}^{n \times 3}, \quad \widehat{\mathcal{M}}_u = \widetilde{\mathcal{M}}_{ED} + \widehat{\Delta \mathbf{u}}_{e \rightarrow u}$$

 Predicted ED displacements (with $\mathbf{H}^{(0)} \leftarrow \widetilde{\mathcal{M}}_u$):

$$\widehat{\Delta \mathbf{u}}_{u \rightarrow e} = \text{MLP}^{u \rightarrow e}(\text{Dropout}(\mathbf{Z}_{\text{Fusion}})) \in \mathbb{R}^{n \times 3}, \quad \widehat{\mathcal{M}}_{ED} = \widetilde{\mathcal{M}}_u + \widehat{\Delta \mathbf{u}}_{u \rightarrow e}$$

Loss Functions and Parameter Update

$$\mathcal{L}_{\text{total}} = \left\| \widehat{\mathcal{M}}_u - \mathcal{M}_u \right\|^2 + \lambda_{\text{cycle}} \cdot \left\| \widehat{\mathcal{M}}_{ED} - \widetilde{\mathcal{M}}_{ED} \right\|^2, \quad \Theta \leftarrow \Theta - \eta \cdot \nabla_{\Theta} \mathcal{L}_{\text{total}}$$

generated 60 distinct ED shapes by drawing unique coefficient sets from this distribution. The resulting surface meshes were then converted into volumetric tetrahedral meshes using Gmsh [117]. Due to meshing variability, the final FEA meshes did not share an identical node structure, node counts ranged from 600 to 1000, and tetrahedral elements from 1800 to 2500. All meshes were visually inspected to ensure anatomical plausibility and sufficient geometric quality for downstream simulation.

After mesh generation, we constructed supervised training pairs $(\mathcal{M}_u, \widetilde{\mathcal{M}}_{ED})$ using a FEA unloading–reloading strategy. The unloaded geometry is first obtained through inverse FEA, and then the loading is reapplied to the unloaded geometries to obtain the ED geometry. The unloaded and ED pair from the reloading exercise is used for training, while the original ED geometry is discarded. This is because the inverse FEA unloading process is typically under-constrained and can have non-unique solution, while the forward FEA loading process deterministically generates the ED geometry. Inverse FEA is also not as accurate due to linearization during the iterative process. Using the unloaded-reloaded pair rather than the ED and unloaded pair for supervision is thus a better choice.

Table 5.1: **Simulation parameters used in the FEA pipeline.** Pressure, fiber orientation, and the Fung stiffness scaling coefficient C were systematically varied, while the remaining dimensionless stiffness ratios b_{ff}, b_{xx}, b_{fx} were fixed across all simulations.

Parameter	Values
<i>Varied across simulations</i>	
EDP	$\{4, 6, 8, 10, 12, 14\}$ mmHg
Fung stiffness scaling coefficient C	$\{50, 100, 150, 200, 250, 300\}$ Pa
Endocardial fiber angle θ_{endo}	$\{60^\circ, 65^\circ, 70^\circ\}$
Epicardial fiber angle θ_{epi}	$\{-60^\circ, -65^\circ, -70^\circ\}$
<i>Fixed across simulations</i>	
Fung anisotropic coefficients	$b_{ff} = 29.9, b_{xx} = 13.3, b_{fx} = 26.6$

To simulate diverse physiological scenarios, we systematically varied key input parameters as detailed in Table 5.1 during the unloading–reloading process. For each of the 60 ED shapes, we evaluated all combinations of the discrete EDP, stiffness and helix-angle values (Cartesian grid), rather than random sampling. EDP was sampled over a range that covers normal LV filling pressures (approximately 5–12 mmHg as reported in [155]) and a small extension into mildly elevated values. The fibre architecture was prescribed using a rule-based transmural helix-angle distribution centred on the canonical pattern from approximately $+60^\circ$ at the endocardium to -60° at the epicardium, with small perturbations to reflect inter-subject variability, in line with histological and Diffusion Tensor (DT)-MRI observations of normal ventricles [156]. The Fung

stiffness scaling coefficient C was varied around literature values commonly used for normal human LV myocardium (on the order of 100 Pa) and extended towards softer (~ 50 Pa) and stiffer (~ 300 Pa) settings that have been adopted in FEA models of hypertrophic or fibrotic ventricles [70, 157]. This combinatorial design resulted in a total of 10,350 successful unloading–reloading simulations after discarding parameter combinations for which the inverse FEA solver did not converge, spanning both anatomical shape and physiological conditions. To evaluate model generalization, we designed two training–testing splits targeting distinct sources of variability.

5

Shape Splits: To evaluate the model’s ability to generalize to unseen anatomical geometries, we split the 60 ED shapes into 42 for training (6,945 samples, 67%) and 18 for testing (3,405 samples, 33%). All global parameter combinations associated with each shape were preserved within the same split to prevent shape leakage. This ensured that test-time geometries were entirely unseen during training, while the model still operated within familiar physiological ranges. This protocol reflected a typical deployment scenario in clinical applications, where patient anatomies were novel but physiological parameters lay within the trained distributions. This shape-based split served as the default evaluation setting and was consistently used across all ablation and comparison experiments unless otherwise noted.

Global Parameter Splits: To evaluate the model’s robustness to unseen physiological conditions, we adopted a LOVO cross-validation strategy for each global parameter ($P, C, \theta_{\text{endo}}, \theta_{\text{epi}}$). In each run, one specific parameter value was held out during training and used exclusively for testing, while all other values were retained. This setup allowed a controlled assessment of the model’s ability to generalize across untrained physiological conditions. All parameter values were normalized based on the training set prior to model input.

5.2.3.2 Implementation Details

Software and hardware: All experiments were implemented in Python using PyTorch and PyTorch Geometric. Training and inference were conducted on a workstation equipped with an NVIDIA RTX 3090 GPU (24 GB VRAM). CUDA (v11.1) and cuDNN optimizations were enabled. *HeartUnloadNet* code is available at <https://github.com/SiyuMU/HeartUnloadNet>. FEA code is available at <https://github.com/WeiXuanChan/heartFEM> [FEniCS, Second-order tetrahedral elements (C3D10)].

Training hyperparameters: The model was trained using the AdamW optimizer with a learning rate of 0.001 and weight decay of 10^{-4} . A total of 1000 training epochs were used with early stopping (patience = 50). Dropout was set to 0.1 in all attention and decoder layers. Mixed-precision training was disabled to ensure numerical stability. The hidden dimension was set to 128, with 4 attention heads used throughout the network. The cycle consistency loss was weighted by 0.2. Global parameters $(P, C, \theta_{\text{endo}}, \theta_{\text{epi}})$ were normalized to $[0, 1]$ before being input into the encoder, and all mesh coordinates were globally normalized to have zero mean and unit scale.

Evaluation metrics: Model performance was evaluated using four geometric metrics between predicted and ground-truth unloaded meshes, as summarized in Table 5.2. It should be noted that the DSC was adapted to a node-level formulation, computed as the percentage of mesh nodes whose Euclidean distance error fell below a 1 mm threshold. This leverages one-to-one correspondence and identical topology to assess overlap directly in mesh space without voxelization, thus preserving spatial resolution and enhancing sensitivity to subtle differences. Such node-level, thresholded overlap metrics have been previously adopted in mesh-based simulations such as [158].

Table 5.2: **Evaluation metrics used for geometric and topological comparison.**

Metric	Description
DSC	Proportion of nodes with error < 1 mm based on one-to-one correspondence
HD	Maximum point-wise deviation between predicted and ground-truth surfaces
Mean Distance (MD)	Mean Euclidean distance between corresponding nodes (global accuracy)
Standard Deviation (SD)	Standard deviation of node-wise distance errors (spatial consistency)

5.3 Experiments and Results

5.3.1 Quantitative Performance Comparison

Table 5.3: **Quantitative performance comparison of baseline methods for predicting unloaded LV geometry from the ED state.**

Method	DSC $\uparrow$	HD $\downarrow$ (cm)	MD $\downarrow$ (cm)	SD $\downarrow$ (cm)	Time $\downarrow$ (s)
GCN	0.674 ± 0.169	0.263 ± 0.073	0.070 ± 0.017	0.041 ± 0.010	0.021 ± 0.002
PointNet++	0.602 ± 0.312	0.216 ± 0.085	0.093 ± 0.048	0.038 ± 0.017	0.018 ± 0.002
Inverse FE	0.811 ± 0.094	0.185 ± 0.080	0.097 ± 0.022	0.099 ± 0.010	7238.1 ± 678.2
<i>HeartUnloadNet</i> (ours)	0.986 ± 0.023	0.083 ± 0.028	0.028 ± 0.010	0.013 ± 0.004	0.023 ± 0.002

Table 5.3 summarized the quantitative performance of our proposed *HeartUnloadNet* against three baseline methods: a GCN[147], PointNet++[159], and a conventional inverse FEA solver that

applied the widely used backward displacement method by Finsberg et al. [160], which iteratively estimated unloaded geometry by simulating inverse FEA unloading and then forward reloading to match target pressure conditions. *HeartUnloadNet* consistently outperformed all baselines across all four geometric metrics, achieving the highest DSC (0.986 ± 0.023) and the lowest HD (0.083 ± 0.028), MD (0.028 ± 0.010), and SD (0.013 ± 0.004).

Notably, GCN outperformed PointNet++ (e.g., GCN 0.674 vs. 0.602), likely due to its use of mesh connectivity, which implicitly encoded structural information such as neighbourhood topology and element volume, important for preserving biomechanical consistency. Building upon this advantage, we adopted the GCN framework as the architectural backbone of *HeartUnloadNet*, progressively integrating pooling, co-attention fusion, and cycle consistency modules to enhance both geometric fidelity and physiological awareness.

Furthermore, compared to the conventional inverse FEA solver, which took over 7,000 seconds per case or roughly 2 hours, *HeartUnloadNet* achieved markedly better geometric accuracy (DSC: 0.986 vs. 0.811) while reducing inference time by more than five orders of magnitude (~ 0.02 s). Notably, the inverse FE method exhibited limited accuracy, with a DSC of 0.81 and HD of 0.185. This was likely due to its first-order approximations when generating unloading deformations, designed to reduce computational burden and improve convergence, but the technique nonetheless required substantial time. As such, despite the FEA method’s strong physical grounding, it had limited practicality for large-scale or real-time applications. In contrast, *HeartUnloadNet* offered near-instantaneous predictions with accuracy comparable to or exceeding that of the inverse FEA baseline, making it far more suitable for clinical deployment or simulation-intensive tasks.

5.3.2 Reduced Supervised Learning Performance

To investigate the capability of the cycle consistency in *HeartUnloadNet* to enable reduced supervision, we conducted a controlled experiment where the SR (i.e., the percentage of labelled training samples) was progressively reduced from full supervision (100%) down to an extremely low partial supervision (1%). For each SR, we randomly split the training set into a labelled subset and an unlabelled subset. Samples in the labelled subset were used with both the supervised displacement loss towards the FEA-unloaded ground truth and the cycle-consistency loss, whereas samples in the unlabelled subset contributed only through the self-supervised cycle-consistency term, without accessing the FEA-unloaded labels. In this sense, only a small fraction of the data provided strong supervision, while the remaining cases supplied a weaker supervisory signal via

Table 5.4: **Performance under varying levels of supervision.** We compared the full model *HeartUnloadNet* (A0) with a variant (A1) that removes the cycle consistency loss. Each model was trained under decreasing SR, from full (100%) to extreme low-supervision (1%).

Model ID	SR	DSC $\uparrow$	HD $\downarrow$ (cm)	MD $\downarrow$ (cm)	SD $\downarrow$ (cm)
A0: <i>HeartUnloadNet</i>	100%	0.986 ± 0.023	0.083 ± 0.028	0.028 ± 0.010	0.013 ± 0.004
	30%	0.982 ± 0.028	0.089 ± 0.031	0.029 ± 0.011	0.014 ± 0.005
	15%	0.977 ± 0.038	0.081 ± 0.030	0.030 ± 0.011	0.013 ± 0.005
	3%	0.971 ± 0.045	0.085 ± 0.031	0.032 ± 0.011	0.014 ± 0.005
	1%	0.619 ± 0.251	0.188 ± 0.045	0.070 ± 0.024	0.033 ± 0.008
A1: A0 w/o cycle loss	100%	0.933 ± 0.143	0.080 ± 0.037	0.033 ± 0.017	0.014 ± 0.007
	30%	0.906 ± 0.126	0.101 ± 0.037	0.039 ± 0.015	0.017 ± 0.008
	15%	0.888 ± 0.118	0.133 ± 0.032	0.043 ± 0.013	0.021 ± 0.006
	3%	0.712 ± 0.210	0.174 ± 0.048	0.062 ± 0.019	0.030 ± 0.008
	1%	0.514 ± 0.197	0.228 ± 0.068	0.084 ± 0.023	0.046 ± 0.015

cyclic reconstruction.

Table 5.4 reported the model’s unloaded-state geometric prediction accuracy under each setting. We compared the fully supervised model (A0), which included the cycle consistency loss, against a variant (A1) that disabled this loss component. Both models were evaluated under identical SR conditions.

As shown in the table, *HeartUnloadNet* (A0) maintained robust performance even under significantly reduced supervision. With just 3% labelled data (around 200 cases), A0 still achieved a DSC of 0.971 ± 0.045 , demonstrating strong generalization with minimal supervision. In contrast, A1, which lacked the cycle loss, suffered a drastic performance decline under the same condition (DSC dropped to 0.712 ± 0.210 , and HD increased from 0.080 to 0.174). This highlighted the critical role of cycle consistency as a form of implicit regularization that enforced geometric coherence and helped prevent overfitting when direct supervision was scarce. These findings suggested that cycle consistency was not merely auxiliary, but essential for learning stable and accurate mappings in reduced supervised deformation modelling.

5.3.3 Ablation Studies

To demonstrate the optimization of *HeartUnloadNet*’s design, we performed ablation studies in this subsection. We first conducted a stepwise ablation starting from the full model (A0), progressively removing or altering key modules in variants A1–A5, as shown in Table 5.5. Next, we carried out a component replacement study, where each of the variants C0–C6 had a key module (encoder, decoder, or pooling) replaced (Table 5.6).

Table 5.5: **Stepwise ablation study of key architectural components in *HeartUnloadNet*.** Models A1–A5 were derived by progressively removing individual modules from the full model A0, highlighting the contribution of each component to overall performance.

Model ID	Components Removed	DSC $\uparrow$	HD $\downarrow$ (cm)	MD $\downarrow$ (cm)	SD $\downarrow$ (cm)
A0	Full <i>HeartUnloadNet</i>	0.986 $\pm$ 0.023	0.083 $\pm$ 0.028	0.028 $\pm$ 0.010	0.013 $\pm$ 0.004
A1	A0 – Cycle Consistency	0.933 $\pm$ 0.143	0.080 $\pm$ 0.037	0.033 $\pm$ 0.017	0.014 $\pm$ 0.007
A2	A1 – Co-attention Fusion	0.930 $\pm$ 0.106	0.094 $\pm$ 0.024	0.045 $\pm$ 0.009	0.015 $\pm$ 0.005
A3	A2 – Average Pooling	0.823 $\pm$ 0.180	0.135 $\pm$ 0.054	0.049 $\pm$ 0.017	0.024 $\pm$ 0.008
A4	A3 – GAT	0.778 $\pm$ 0.165	0.163 $\pm$ 0.035	0.053 $\pm$ 0.015	0.027 $\pm$ 0.006
A5	A4 – Dual Input (GCN base model only)	0.674 $\pm$ 0.169	0.263 $\pm$ 0.073	0.070 $\pm$ 0.017	0.041 $\pm$ 0.010

Table 5.6: **Component-wise Ablation Study of key architectural components in *HeartUnloadNet*.** Models C0–C6 replaced encoder, decoder, or pooling modules.

Model ID	Replaced	Alternative	DSC $\uparrow$	HD $\downarrow$ (cm)	MD $\downarrow$ (cm)	SD $\downarrow$ (cm)
A0	–	–	0.986 $\pm$ 0.023	0.083 $\pm$ 0.028	0.028 $\pm$ 0.010	0.013 $\pm$ 0.004
C0	GAT Encoder	GATv2 Encoder	0.985 $\pm$ 0.022	0.074 $\pm$ 0.023	0.028 $\pm$ 0.009	0.013 $\pm$ 0.005
C1	GAT Encoder	GATv2 + 3% SR	0.600 $\pm$ 0.246	0.177 $\pm$ 0.043	0.074 $\pm$ 0.025	0.034 $\pm$ 0.009
C2	GAT Encoder	GraphSAGE Encoder	0.963 $\pm$ 0.056	0.084 $\pm$ 0.030	0.032 $\pm$ 0.012	0.015 $\pm$ 0.006
C3	GAT + Pooling	MeshGraphNet	0.942 $\pm$ 0.029	0.088 $\pm$ 0.029	0.030 $\pm$ 0.011	0.017 $\pm$ 0.005
C4	GAT Encoder	Graph Transformer	0.911 $\pm$ 0.124	0.095 $\pm$ 0.029	0.039 $\pm$ 0.015	0.018 $\pm$ 0.006
C5	MLP Decoder	GAT Decoder	0.725 $\pm$ 0.204	0.226 $\pm$ 0.058	0.060 $\pm$ 0.020	0.034 $\pm$ 0.008
C6	Average Pooling	Max Pooling	0.934 $\pm$ 0.111	0.089 $\pm$ 0.028	0.035 $\pm$ 0.014	0.016 $\pm$ 0.006

In the progressive ablation study (Table 5.5), the cyclic consistency loss was removed from A1, resulting in a clear performance degradation (DSC dropped from 0.986 to 0.933) and a pronounced increase in variability (± 0.143), suggesting its critical role in enforcing geometric coherence and stabilizing training. HD showed a slight but negligible improvement. In A2, replacing the co-attention fusion with naive feature concatenation deteriorated HD and MD with minimal effects on DSC, underscoring the necessity of attention-based feature fusion for stable and accurate deformation prediction.

In A3, removing the pooling mechanism eliminated global mesh context, leading to a substantial reduction in DSC and an increase in HD, which highlighted the importance of hierarchical global feature aggregation for capturing whole LV geometry. In A4, replacing the GAT encoder with a standard GCN resulted in a further substantial decline in performance across all metrics, underscoring the superiority of attention-based neighbourhood modelling in capturing important local features. Finally, A5 removed the dual-input structure and simply broadcast global physiological parameters to each node. This variant yielded the poorest performance (DSC: 0.674 ± 0.169), underlining the necessity of disentangled and context-aware fusion between mesh geometry and physiological priors for accurate biomechanical inference.

In the component replacement study (Table 5.6), C0 replaced the original GAT encoder with GATv2 [161], and consequently achieved slightly reduced DSC accuracy (DSC: 0.985 vs. 0.986) and slightly better HD (0.074 vs. 0.083). This suggested that GAT and GATv2 performed similarly

when sufficient supervision was available. However, in C1, using the same GATv2 under only 3% supervision led to a major drop in performance (DSC: 0.600 ± 0.246). This was likely because GATv2 computed attention weights using both source and target node features through a learnable function, making it more flexible but also more prone to overfitting when data was scarce. In contrast, the original GAT used a shared attention mechanism that offered stronger inductive bias and greater stability under reduced supervision. These findings reaffirmed that high-capacity architectures like GATv2 were less suitable in low-data regimes, as their flexibility increased the risk of overfitting despite regularization.

Additionally, replacing the encoder with GraphSAGE [138] (C2) or MeshGraphNet [137] (C3) led to a noticeable decline in overall performance. These methods, while effective in other graph domains, lacked the fine-grained attention mechanism of GAT that may have been important for capturing irregular mesh geometry and directional anisotropy in biomechanical structures. Similarly, substituting the encoder with a graph transformer [162] (C4) resulted in further degradation (DSC: 0.911), likely due to the absence of localized inductive bias and the model's reliance on large datasets for effective generalization. In contrast, GAT's edge-aware attention enabled adaptive weighting of neighbouring nodes based on geometric context, making it particularly well-suited for spatially structured data such as tetrahedral heart meshes.

Moreover, in C5, replacing the original MLP decoder with a GAT-based decoder led to substantial degradation across all metrics (e.g., SD increased from 0.013 to 0.034, DSC dropped to 0.725). This suggested that additional message passing during decoding was not only unnecessary but potentially detrimental. Since the decoder operated on already fused, spatially structured features, further aggregation likely introduced redundant interactions and amplified local noise. The higher parameter complexity of GAT may also have destabilized training, especially under biomechanical constraints that favoured smooth, globally coherent deformations. In C6, substituting average pooling with max pooling resulted in moderate degradation, indicating that average pooling better preserved subtle global geometry.

These findings highlighted the optimality of the proposed *HeartUnloadNet* configurations.

5.3.4 Generalization to Unseen Global Physiological Values

Table 5.7 reported the results of LOVO validation across global physiological parameters, including pressure P , stiffness C , and fiber orientations $\theta_{\text{endo}}, \theta_{\text{epi}}$. Despite the exclusion of specific values during training, *HeartUnloadNet* consistently achieved high accuracy on the held-out cases, with

Table 5.7: **LOVO cross-validation results.** Each variable group (pressure, stiffness, fiber orientation) is excluded during training to evaluate generalization.

Parameter Group	Test Case	DSC $\uparrow$	HD $\downarrow$ (cm)	MD $\downarrow$ (cm)
Pressure (P)	P (4 mmHg)	0.999	0.007	0.003
	P (6 mmHg)	0.999	0.008	0.003
	P (10 mmHg)	0.999	0.007	0.002
Stiffness (C)	C (50)	0.942	0.011	0.005
	C (150)	1.000	0.006	0.002
	C (300)	1.000	0.006	0.002
Endo Helix (θ_{endo})	Endo (60°)	0.998	0.008	0.003
	Endo (65°)	0.999	0.007	0.003
	Endo (70°)	0.999	0.008	0.003
Epi Helix (θ_{epi})	Epi (60°)	1.000	0.006	0.002
	Epi (65°)	1.000	0.006	0.002
	Epi (70°)	0.999	0.008	0.003

DSC exceeding 0.998 and HD below 0.01 cm in most settings. The only drop appeared at $C = 50$ (DSC: 0.942), suggesting a mild sensitivity to low-stiffness outliers and the importance of having sufficient training cases at outlying stiffnesses. Overall, the results demonstrated strong interpolation capability within the physiological range.

5.3.5 Qualitative Visualization of Predictions

Figure 6.12 presented heatmaps of vertex-wise displacement errors for two representative test cases from the test set. Given that the overall prediction error was on the sub-centimetre scale, direct visual comparison of mesh displacements was not intuitive. Therefore, we visualized the error distribution using colormaps rendered on the ED mesh to highlight subtle discrepancies. Warmer colours indicated larger local deviations, and all methods were compared under a unified scale and view orientation.

In Case 1, the inverse FEA method exhibited globally elevated errors (reddish colour), indicating a general mismatch between the predicted and true unloaded shapes. This was evident in both views, where the inverse FE result showed notable shape distortion compared to the ground truth. For GCN and PointNet++, the errors were primarily concentrated around the apex and basal regions. In Case 2, PointNet++ performed poorly, with visibly misaligned predictions near the apex. GCN also showed moderate inconsistencies along the myocardial wall. In contrast, both inverse FEA and *HeartUnloadNet* performed better in Case 2, though *HeartUnloadNet* maintained significantly lower errors (mostly under 0.1 cm) in both cases, demonstrating clear superiority in both spatial precision and robustness.

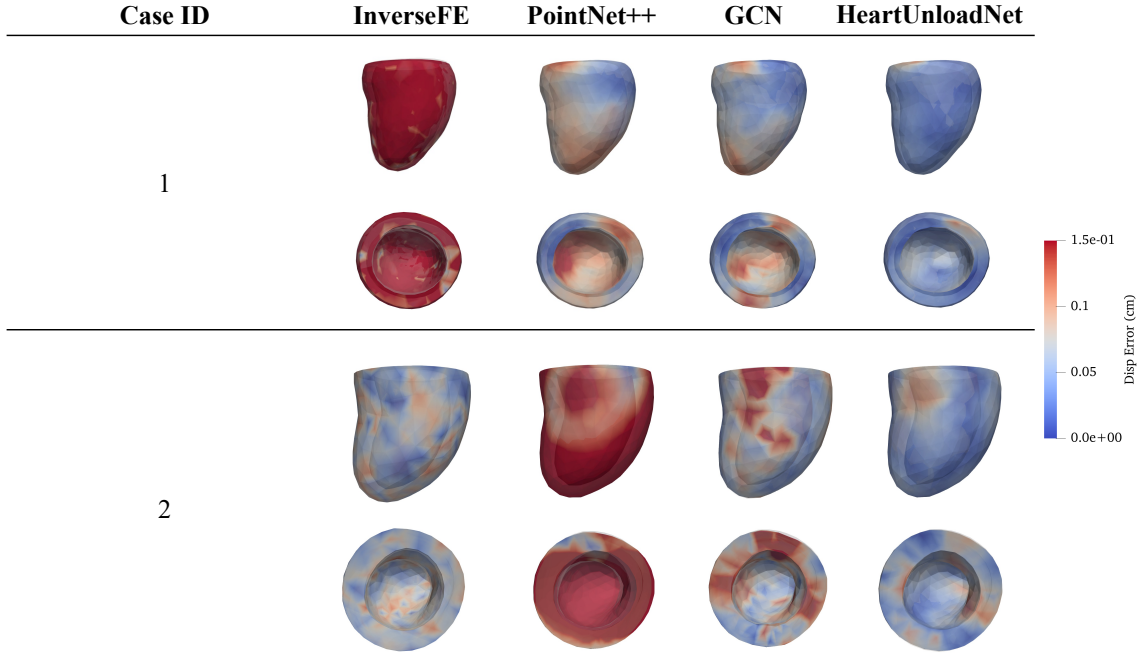


Figure 5.3: **Qualitative heatmap visualization of vertex-wise prediction errors on two representative test cases.** Each row shows a different method (Inverse FEA, PointNet++, GCN and *HeartUnloadNet*), and each column corresponds to one test case. Warmer colors indicate higher local displacement errors.

5.3.6 Advantages over PCA-Based and Inverse FEA Approaches

Table 5.3 showed that *HeartUnloadNet* reduced inference time by over $10^5 \times$ while improving accuracy compared to the traditional inverse FEA solver based on the Finsberg inverse step method [160]. Although inverse FEA can be accelerated using reduced-order models such as PCA-based statistical shape models [142], these approaches remained limited in capturing complex unloading deformations.

To enable a fair comparison, we constructed a PCA model using displacement fields (i.e., node-wise differences between ED and unloaded meshes) across the training set. For each test case, we predicted the deformation by regressing ED-derived PCA coefficients and reconstructing the corresponding displacement field. The PCA-based predictions resulted in a HD of 1.02 ± 0.07 cm. In contrast, *HeartUnloadNet* achieved significantly lower errors, with a HD of 0.083 ± 0.028 cm.

Figure 5.4 visualized the deformation heatmaps predicted by PCA and *HeartUnloadNet*. Ground truth vectors were omitted for clarity, as the predictions from *HeartUnloadNet* were nearly indistinguishable from them. In contrast, PCA predictions exhibited clear discrepancies, highlighting its limitations in capturing subtle, spatially distributed deformations using a low-dimensional linear

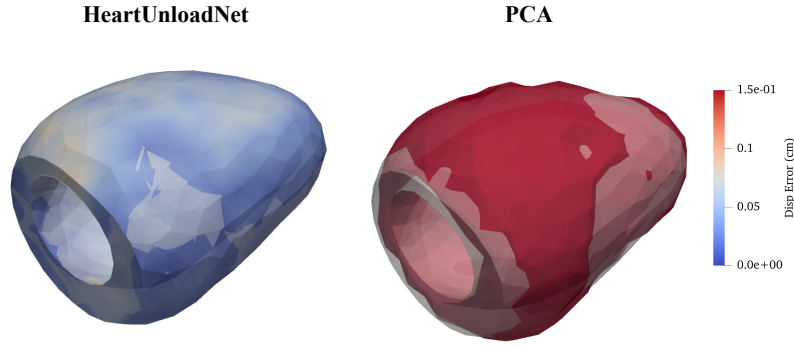


Figure 5.4: **Qualitative heatmap visualization comparing vertex-wise prediction errors of PCA and *HeartUnloadNet* for a representative test case.** Compared to the ground truth, PCA exhibited a HD of 1.02 ± 0.07 cm, whereas *HeartUnloadNet* achieved 0.083 ± 0.028 cm.

basis.

These observations highlighted the inherent limitations of PCA in modeling nonlinear tissue mechanics, where complex deformations could not be adequately represented by a small number of dominant modes. By learning a direct nonlinear mapping from ED meshes and physiological parameters to displacement fields, *HeartUnloadNet* achieved not only higher accuracy and faster inference, but also better preservation of local biomechanical fidelity.

5.4 Discussion

5.4.1 Novelty and Contribution

To the best of our knowledge, this is the first study to address the problem of predicting the unloaded cardiac geometry directly from clinical diastolic states using a graph-based DL framework. No prior literature has proposed a comparable approach, underscoring the novelty of our contribution. In the absence of directly related work, we conducted comprehensive comparisons against a wide range of baselines, including traditional inverse FEA solvers, PCA-based reduced-order models, and DL architectures such as GCN and PointNet++.

The experimental results validated not only the accuracy and efficiency of *HeartUnloadNet*, but also demonstrated its potential to fundamentally shift the approach to unloaded geometry estimation. Unlike conventional inverse FEA pipelines, which rely on iterative solvers, are sensitive to initialization and boundary condition settings, and converge slowly, our framework predicts the unloaded state through a single forward pass. *HeartUnloadNet* consistently outperformed

both Finsberg et al.’s inverse FE [160] and PCA-based reduced-order approaches [142] across all geometric metrics, while reducing inference time by over five orders of magnitude.

Furthermore, extensive ablation studies confirmed the critical role of cycle-consistent training and physiological conditioning in achieving superior performance. These findings suggest that DL, when properly integrated with mesh topology and physiological priors, is not merely a surrogate for inverse finite element analysis, but instead offers a robust and generalizable alternative for tackling biomechanical inverse problems.

5.4.2 Implications

From a biomechanics perspective, accurately estimating the unloaded cardiac geometry is fundamentally important, as it serves as the reference configuration in FEA simulations of cardiac function. In applications such as heart disease analysis or interventional planning, all deformations, stresses, and energy quantities are derived from this baseline. Errors in the unloaded state propagate through downstream simulations, potentially distorting disease characterization, stress quantification, and virtual treatment assessments. A reliable unloaded geometry is therefore essential not only for passive mechanical analysis, but also for patient-specific inverse modelling and therapy optimization.

From a clinical perspective, fast and accurate estimation of the unloaded geometry can also support real-time, image-based cardiac analysis that may be developed in future. *HeartUnloadNet* delivers near-instantaneous predictions with high geometric fidelity, making it well suited for integration into real-time diagnostic pipelines. As cardiac function assessment increasingly shifts toward image-driven, data-centric workflows, efficient estimation of the reference unloaded state will be a critical enabler for real-time stress or contractility evaluation.

5.4.3 Architectural Strengths and Outlook

A key advantage of our proposed framework lies in its mesh-agnostic design, which eliminates the need for a fixed nodal configuration or template mesh. Unlike many prior approaches that required point correspondences or mesh registration to enable learning, our method directly operated on arbitrary tetrahedral meshes, leveraging the inherent topology through GAT-based message passing. This flexibility allowed the model to accommodate diverse cardiac geometries, making it readily applicable to patient-specific data without labor-intensive preprocessing.

In addition, our architectural choices, including the use of graph attention, co-attentive global fusion, and cycle-consistent supervision, collectively contributed to the model’s superior prediction accuracy across all baselines. Beyond demonstrating performance, our study offered a comprehensive architectural evaluation, ablation and component-replacement experiments. Our results can provide insight into which architectural strategies are most effective for future model design in data-driven biomechanics.

Finally, while the present implementation instantiates the passive myocardium with a transversely anisotropic Fung-type law (fixed b_{ij} and varying global stiffness C), *HeartUnloadNet* itself is agnostic to the specific choice of constitutive model. In principle, the same training pipeline could be re-run with alternative myocardium laws such as Holzapfel–Ogden [62] by regenerating the FE unloading dataset under different parameter sets, with no change to the network architecture or loss design.

We believe that *HeartUnloadNet* offers a promising foundation for future development of end-to-end cardiac mechanics predictors. Although the current work focused on diastolic unloading, our GAT-based architecture can easily be extended to model dynamics during systole or across the cardiac cycles. Such future work might benefit further from incorporating physics-informed constraints, or residual formulations to jointly learn loading and unloading trajectories in a physiologically accurate, interpretable manner.

5.4.4 Limitations and Future Work

HeartUnloadNet demonstrated strong performance on data representing healthy anatomical variations, suggesting that the proposed architecture can robustly learn the ED-to-unloaded mapping over a wide range of normal-like geometries and global stiffness levels. Its generalizability to more diverse pathological cases, however, remains to be systematically evaluated, and future work will investigate its robustness across a broader spectrum of disease conditions.

From a modelling perspective, the present implementation adopts a Fung-type transversely anisotropic passive material law with a fixed set of dimensionless anisotropy parameters b_{ij} , while only the global stiffness scaling parameter C is varied in the synthetic unloading simulations. In addition, and in line with most diastolic patient-specific LV FEA studies, our synthetic unloading dataset assumes spatially homogeneous passive material properties and fibre architecture across the LV, without explicitly modelling regional heterogeneity in stiffness, anisotropy or scar tissue. The

current model should therefore be regarded as calibrated under a commonly used global material parametrisation, rather than as a fully general solution for arbitrarily heterogeneous material fields.

Furthermore, the synthetic training data are rooted in geometries and fibre architectures derived from 1,991 clinically observed ED shapes and in global stiffness values C sampled over a wide range (50–300), which together span a physiologically meaningful spectrum of LV size, shape and passive stiffness around normal and mildly to moderately stiff ventricles. Nonetheless, the present dataset is not designed to exhaustively cover all conceivable extreme pathological remodelling patterns. A single universal model that remains accurate across all possible disease phenotypes would, in principle, require training over a vastly broader parameter space and is unlikely to be practically achievable, instead, we envisage the current *HeartUnloadNet* as a backbone that can be adapted to specific pathological cohorts via transfer learning or fine-tuning once disease-specific FEA or clinical datasets become available.

Future extensions of *HeartUnloadNet* will therefore require regionally heterogeneous FEA datasets, in which stiffness, anisotropy, fibre angles and scar patterns are systematically varied, and the inclusion of suitable material parameters (C, b_{ij}) together with region-wise descriptors as additional network inputs, in order to better capture both global and local remodelling in diseased hearts. In addition, the same graph-based framework can be naturally extended from isolated LV models to biventricular geometries with region labels (LV, RV and septum) and chamber-specific loading and stiffness parameters, enabling applications to conditions such as pulmonary arterial hypertension with RV pressure overload and septal remodelling.

Finally, our FEA training meshes had a moderate resolution (600–1000 nodes, 1800–2500 tetrahedra), which reflects a practical compromise between anatomical fidelity and the need to generate over 10,000 unloading–reloading simulations within reasonable computational time. From the network side, however, *HeartUnloadNet* itself is not tied to a particular mesh density: the GAT encoder–decoder has about 9 MB of learnable parameters, independent of the number of nodes, and the per-iteration cost scales approximately linearly with the node and edge counts. Building on this, future work will focus on generating higher-resolution FEA unloading datasets and applying *HeartUnloadNet* to denser LV meshes (e.g. with $> 10,000$ nodes). Given the modest model size and near-linear scaling, such extensions should remain well within the capability of a single modern GPU (e.g. RTX 3090) after adjusting the batch size, without requiring changes to the core architecture.

5.5 Conclusion

This work introduced a data-driven framework for estimating unloaded LV geometry without relying on traditional physics-based optimization. By leveraging graph attention mechanisms and incorporating biophysical priors (pressure, stiffness, and fiber orientation), our model directly predicted deformation from ED meshes with high geometric fidelity. It eliminated the need for iterative FEA solvers or reduced-order assumptions. On a held-out test set of 6,810 synthetic LV shapes, the network achieved a DSC of 0.986 ± 0.023 and a HD of 0.083 ± 0.028 cm, with an inference time of just 0.023 seconds per case, surpassing inverse FEA in both accuracy and efficiency.

HeartUnloadNet also maintained strong performance under reduced supervision, achieving over 97% DSC with only 200 labeled samples. Its topology-agnostic design supported generalization across varying anatomical and material conditions, highlighting its potential for scalable application in both research and clinical settings. As the first learning-based method to infer unloaded cardiac geometry using explicit physiological priors, this work establishes a foundation for real-time, data-driven cardiac mechanics.

Graph Neural Network Surrogates for Left Ventricular Mechanics over the Cardiac Cycle

Contents

6.1	Introduction	98
6.1.1	Contributions	99
6.2	Methods	100
6.2.1	Problem Domain	100
6.2.2	CardioGraphFENet	102
6.2.3	Lumped parameter model coupling and P-V-t lookup table	110
6.2.4	Data Generation	111
6.2.5	Experimental Setup	113
6.3	Results	113
6.3.1	Overview	113
6.3.2	Baseline Comparison	114
6.3.3	Ablation Study	116
6.3.4	Reduced supervision capability	119
6.3.5	Windkessel-coupled P-V loop evaluation	120
6.4	Discussion	121
6.4.1	Overall performance and physiological fidelity	121
6.4.2	Comparison with existing cardiac surrogate literature	123
6.4.3	Limitations and future work	124
6.5	Conclusion	125

This chapter presents CardioGraphFENet (*CGFENet*) as a GNN surrogate for LV mechanics over the cardiac cycle. The model takes a variable tetrahedral LV mesh together with a target

cavity volume and time as inputs, and predicts the corresponding myocardial displacement field and chamber pressure. In contrast to physics informed approaches such as *IMC-PINN-FE* in Chapter 4, *CGFENet* is fully data driven and does not rely on reduced order modes, which improves geometric flexibility and strengthens generalisation across anatomies. The surrogate follows a volume driven formulation that supports direct evaluation at arbitrary target volumes and enables efficient construction of case level P-V-t maps. To handle heterogeneous meshes, the network combines an attention based graph encoder with global local feature fusion and a recurrent temporal module to capture cycle coherent pressure dynamics. An enhanced cycle consistent bidirectional formulation, building on *HeartUnloadNet* in Chapter 5, is introduced to jointly model loading and inverse unloading within a single network, and to reduce the amount of required FEA supervision with only minor loss in accuracy. Experiments on a large scale FEA dataset of full cycle LV states show that the proposed surrogate improves accuracy over representative graph baselines while producing complete cycle predictions within one minute per case. When coupled with a lumped parameter model, the predicted pressures produce physiologically plausible closed P-V loops and accurate global functional indices across diverse anatomies.

6.1 Introduction

This chapter extends *HeartUnloadNet* in Chapter 5 from an inverse unloading network to a unified full-cycle surrogate for LV mechanics [24]. The objective is near-real-time emulation of cycle-wide FEA outputs on variable tetrahedral meshes, including myocardial deformation and chamber pressure, while retaining inverse capability to recover the zero-pressure reference state.

A key motivation is to overcome limitations of physics informed learning for realistic ventricular mechanics. In Chapter 4, *IMC-PINN-FE* incorporates physical constraints and motion consistency to improve physiological plausibility [23]. However, PINNs still scale poorly to complex LV geometries and high-resolution meshes, and optimisation is sensitive to geometry and mesh quality [21–23]. Reduced-order representations can make optimisation tractable by lowering dimensionality and simplifying geometry, but they also constrain geometric coverage and can struggle to represent rare or previously unseen anatomies, which can be important for novel disease phenotypes. Moreover, many PINN approaches remain patient specific and require per-case training or fine tuning, which limits practical pre-training and hinders cohort-scale near-real-time deployment [22, 23]. These factors motivate a complementary supervised strategy based on large simulation datasets, enabling stable training and fast inference on full-resolution meshes.

Graph based surrogates are well suited to this strategy because they operate directly on unstructured tetrahedral meshes and generalise across varying connectivities and node counts [18–20]. Prior GNN emulators have demonstrated strong accuracy and efficiency for passive ventricular inflation and related restricted settings, and *HeartUnloadNet* further showed that inverse recovery of unloaded geometry can be learned on heterogeneous meshes [18–20, 24]. However, existing models do not yet provide a unified treatment of full-cycle mechanics, where deformation and pressure must remain temporally coherent across both diastole and systole, and where forward loading and inverse unloading are both required for end-to-end case analysis [18, 19, 24].

To address this gap, this chapter presents CardioGraphFENet (*CGFENet*) as a unified GNN surrogate for LV mechanics over the cardiac cycle. The model generalises the bidirectional design of *HeartUnloadNet* from unloading alone to full-cycle prediction, and jointly outputs the displacement field and chamber pressure from a volume and time query on a given tetrahedral mesh [24]. A volume driven formulation enables direct evaluation at arbitrary target volumes and supports efficient construction of case level volume time pressure maps without relying on fixed temporal sampling. Architecturally, *CGFENet* combines an attention based mesh encoder with global local feature fusion to represent heterogeneous meshes, together with a recurrent temporal module to model cycle coherent pressure dynamics. The cycle consistent bidirectional formulation is further strengthened to support both forward loading and inverse unloading within a single network, and to reduce the amount of required FEA supervision with limited loss in accuracy. Finally, the predicted pressure trajectory is coupled to a lumped parameter model to generate physiologically plausible closed pressure volume loops that remain consistent with the predicted mechanics.

6.1.1 Contributions

Our main contributions are fourfold:

1. **Unified loading and unloading within one surrogate:** We propose a cycle-consistent formulation that performs forward prediction of chamber pressure and full-field displacement from prescribed volume and time inputs, and inverse recovery of the zero-pressure reference configuration within a single GNN model.
2. **Full-cycle mechanics on arbitrary tetrahedral meshes:** We enable mesh-agnostic full-cycle prediction of pressure and displacement on diverse LV geometries for any queried cavity volume and time, covering the complete diastolic to systolic loop without mesh registration or reduced-order representations.

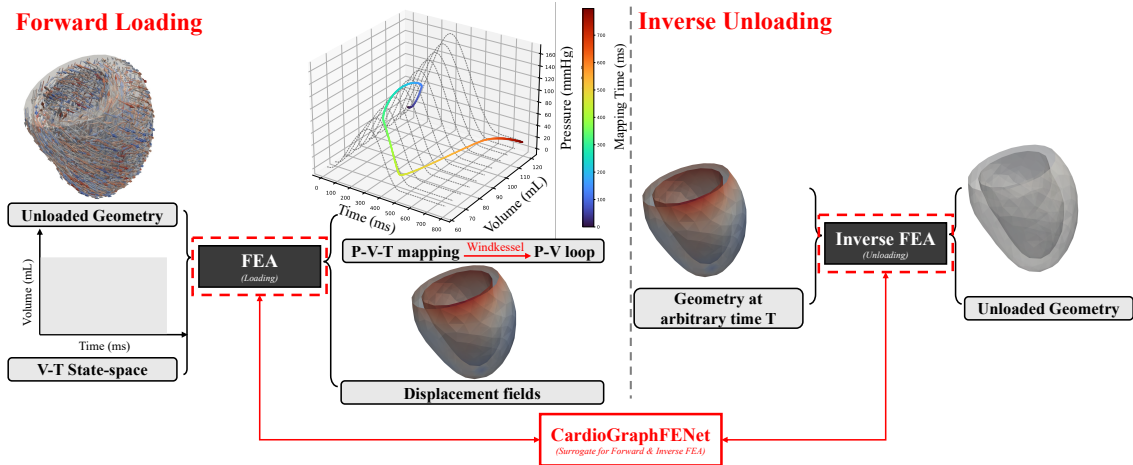


Figure 6.1: **Overview of the problem domain and the role of *CGFENet*:** forward loading maps prescribed (V, t) and unloaded geometry Ω_0 to pressure and displacement. Inverse unloading recovers Ω_0 from a loaded geometry. *CGFENet* replaces both expensive FEA operators with a unified cycle-consistent surrogate.

3. Physiological pressure volume reconstruction via lumped parameter coupling:

We couple the predicted pressure trajectories to a Windkessel model to reconstruct physiologically plausible closed pressure volume loops that remain consistent with the predicted deformation fields.

4. **Large-scale full-cycle LV mechanics dataset for supervised learning:** We create a dataset of 1,814,146 mechanically consistent LV states from 67 anatomies, spanning wide ranges of cavity volume and full-cycle time, and providing pressure, volume, time, and full-field nodal displacements for supervised training and evaluation.

6.2 Methods

6.2.1 Problem Domain

In LV FEA, the cavity pressure P_t , cavity volume V_t , and cardiac time t constitute the key loading quantities governing the ventricular mechanical state. Given the unloaded reference configuration Ω_0 , these variables are mechanically coupled through the constitutive law and boundary conditions, such that prescribing any two of (P_t, V_t, t) implicitly determines the third.

Most existing surrogate FEA approaches, including the PINN frameworks of Buoso et al. [22] and Motiwale et al. [21], employ a pressure-driven formulation. In these methods, the target pressure P_t together with an activation signal, typically a time-dependent active tension $T_{\text{active}}(t)$, is

prescribed, and the equilibrium solution yields the deformation field $\mathbf{u}_t$, from which the cavity volume V_t is computed. Besides pressure-driven analysis, volume-driven simulation is also widely used in practice and is often more compatible with imaging workflows, since ventricular V-t curves can be directly extracted from medical imaging [23, 128]. However, volume-driven FEA is commonly implemented by iteratively running pressure-driven solves and updating P_t until the simulated volume matches the prescribed target, which makes it substantially more expensive than a single pressure-driven analysis. This cost profile makes near-real-time DL surrogates particularly more attractive for volume-driven use cases. Therefore, consistent with *IMC-PINN-FE* [23], this chapter still focuses on the volume-driven problem and learns a direct mapping from (V_t, t) to pressure and full-field displacement $\mathbf{u}_t$.

To achieve this, we adopt a V-t state-space formulation and train a surrogate that directly maps the query pair (V_t, t) to the corresponding chamber pressure and full-field displacement. By sampling a physiologically plausible range of V_t and t during training, the model learns a dense representation of the mechanical response over the cardiac cycle. At inference, the surrogate can be driven by subject-specific V-t curves extracted from medical imaging, or coupled with a lumped parameter circulation model, to generate the complete P-V-t trajectory across the cardiac cycle.

Beyond forward loading, LV modelling also requires solving the inverse FEA problem of recovering the zero-pressure reference configuration Ω_0 from a given loaded geometry Ω_t^{given} . Conventional unloading is computationally expensive because it relies on repeated updates of Ω_0 and repeated forward solves until the simulated deformation and cavity volume match the observed state. This chapter therefore integrates the inverse unloading functionality developed in *HeartUnloadNet* in Chapter 5 into the same full-cycle surrogate framework.

Thus, LV biomechanics involves two complementary operators as Fig. 6.1 shown:

$$\begin{aligned} \text{Forward loading } (\mathcal{F}_{\text{load}}) : \quad & (V_t, t, \Omega_0) \rightarrow \{P_t, \mathbf{u}_t\}, \\ \text{Inverse unloading: } (\mathcal{F}_{\text{unload}}) : \quad & (\Omega_t^{\text{given}}, t) \rightarrow \Omega_0. \end{aligned}$$

The aim of this work is to replace both high-cost FEA operators with a unified, cycle-consistent surrogate capable of performing forward loading and zero-pressure unloading on arbitrary LV meshes.

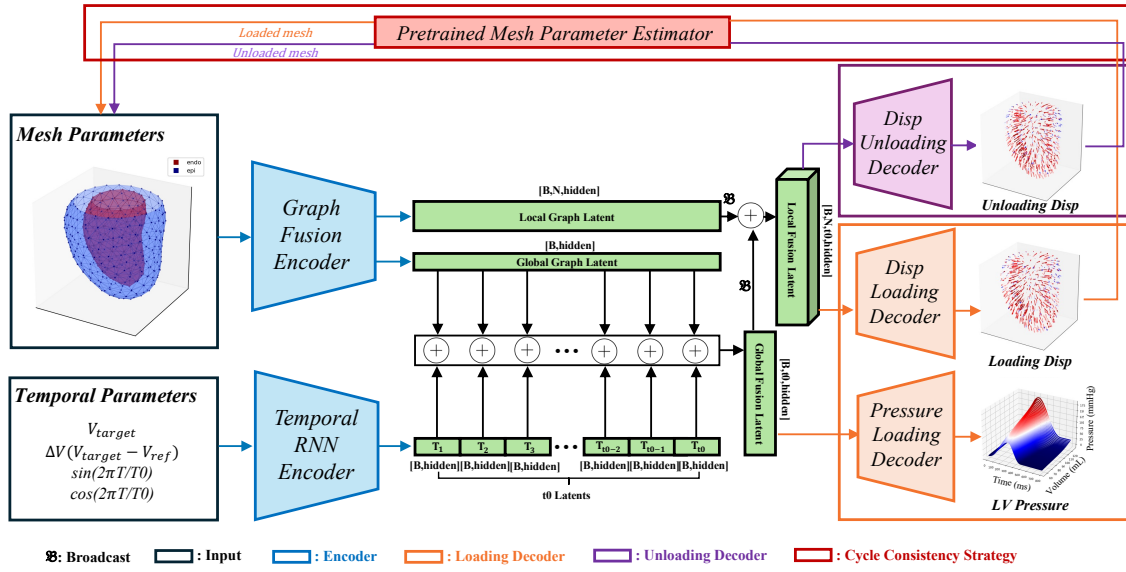


Figure 6.2: **Overall architecture of *CGFENet***: A dual-stream design encodes LV geometry (graph encoder) and the V-t signal (temporal encoder) into a shared latent space. The fused latent drives two task heads for forward loading (pressure and displacement) and inverse unloading (zero-pressure recovery), while a cycle-consistent constraint regularises the two mappings across the cardiac cycle. A pre-trained mesh-parameter estimator (frozen) provides the shared global features used in Fig. 6.3.

6.2.2 CardioGraphFENet

6.2.2.1 Overview

Fig. 6.2 presents the overall schematic of *CGFENet*, a unified architecture developed to approximate both forward loading and inverse unloading FEA prediction. The network receives two groups of inputs, geometric mesh parameters and temporal loading parameters, which are processed by two dedicated encoders. The mesh branch employs a graph based encoder to extract local and global structural features from the ventricular geometry, whereas the temporal branch uses an RNN based encoder to characterise the evolution of volume and time dependent loading signals. Their latent representations are subsequently fused in the central module (green blocks in Fig. 6.2), after which two task specific decoders are applied. The loading decoder outputs the predicted displacement field and ventricular pressure, while the unloading decoder estimates the displacement required to recover the zero-pressure configuration.

In the following subsections, we detail each component of the architecture, including the graph fusion encoder, temporal RNN encoder, latent fusion mechanism, cycle consistency strategy, and the inference procedure.

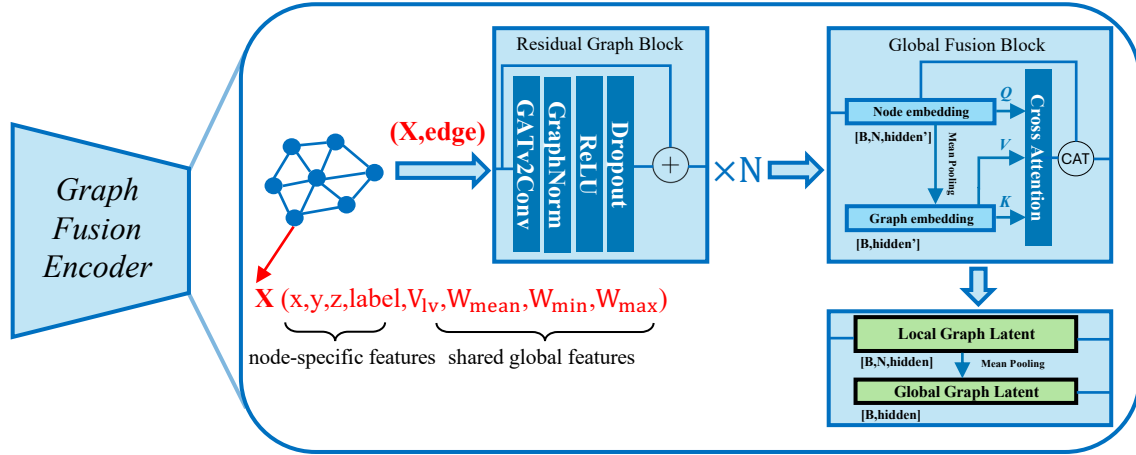


Figure 6.3: **Schematic of the Graph Fusion Encoder:** Node features include both node-specific attributes and shared global descriptors, and are processed through stacked GATv2 residual blocks. A global graph descriptor is obtained by mean pooling and fused back into node embeddings via cross-attention, enabling the encoder to integrate local mesh connectivity with chamber level mechanical context.

6.2.2.2 Graph Fusion Encoder

The Graph Fusion Encoder (Fig. 6.3) operates on the geometric graph $(\mathbf{X}, \text{edge})$, where $\mathbf{X} \in \mathbb{R}^{B \times N \times 8}$ stores an 8-dimensional feature vector for each of the N nodes in a batch of B meshes. Each node feature $\mathbf{X}_i$ contains (i) node-specific attributes $(x_i, y_i, z_i, \text{label}_i)$, where (x_i, y_i, z_i) denote the spatial coordinates and $\text{label}_i \in \{0, 1, 2\}$ identifies endocardial, epicardial, or interior nodes, and (ii) four global descriptors $(V_{LV}, W_{\text{mean}}, W_{\text{min}}, W_{\text{max}})$ broadcast to all nodes, characterising the instantaneous cavity volume and wall thickness distribution. These descriptors correspond to quantities used in FEA formulations of cardiac mechanics, where coordinates determine deformation gradients, anatomical labels impose pressure-loaded and traction-free boundary conditions, cavity volume governs the P-V coupling relation, and wall thickness measures capture structural heterogeneity that influences regional stiffness and deformation. Incorporating these FEA relevant geometric quantities provides an inductive bias consistent with the governing PDEs and supports latent representations that more faithfully reflect LV mechanical behaviour.

We initialise the node embedding as $\mathbf{h}_i^{(0)} = \mathbf{X}_i$ and process the graph using N residual blocks based on GATv2 [161]. Compared with GCN [147] and GraphSAGE [138], GATv2 is better suited to LV meshes, as confirmed by our ablation experiments. This is mainly because LV meshes exhibit irregular connectivity and pronounced regional and directional heterogeneity related to myocardial fibre architecture. Attention based message passing assigns state aware and direction aware weights to neighbours, allowing the network to capture spatially varying mechanical patterns

more effectively than shared weight or distance based aggregation.

For node i at layer ℓ , multi-head attention with M heads is computed as:

$$e_{ij}^{(m)} = (\mathbf{a}^{(m)})^\top \phi\left(W^{(m)}[\mathbf{h}_i^{(\ell)} \parallel \mathbf{h}_j^{(\ell)}]\right), \quad (6.1a)$$

$$\alpha_{ij}^{(m)} = \frac{\exp(e_{ij}^{(m)})}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik}^{(m)})}, \quad (6.1b)$$

$$\mathbf{h}_{i,\text{attn}}^{(\ell)} = \parallel_{m=1}^M \sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(m)} W^{(m)} \mathbf{h}_j^{(\ell)}. \quad (6.1c)$$

where $\mathcal{N}(i)$ denotes the neighbours of node i , and $W^{(m)}$ and $\mathbf{a}^{(m)}$ are projection matrix and attention vectors for head m . The operators $[\parallel]$ and $\phi(\cdot)$ denote concatenation and an activation function. We adopt LeakyReLU for ϕ , and use $M = 4$ attention heads in our implementation.

Each GATv2 layer is followed by a residual refinement to improve optimisation stability and representational capacity:

$$\mathbf{h}_i^{(\ell+1)} = \mathbf{h}_i^{(\ell)} + \text{Dropout}\left(\text{ReLU}\left(\text{GraphNorm}\left(\mathbf{h}_{i,\text{attn}}^{(\ell)}\right)\right)\right). \quad (6.2)$$

Beyond local message passing, LV mechanics include chamber level constraints arising from the FEA weak form, such as global deformation modes, near incompressible volume behaviour, and pressure volume coupling, which cannot be captured by neighbourhood aggregation alone. Similar to *HeartUnloadNet*, to incorporate these global effects, the encoder employs an explicit global local fusion mechanism. A global graph descriptor is first obtained through mean pooling [163] over node embeddings, and this descriptor is then injected back into each node through a cross attention operation [151], in which node embeddings act as queries and the global descriptor provides keys and values.

$$g_{\text{glob}} = \text{MeanPool}\left(\mathbf{h}^{(\ell+1)}\right) \in \mathbb{R}^{B \times \text{hidden}'}, \quad (6.3a)$$

$$Z_{\text{local}} = \text{softmax}\left(\frac{QK^\top}{\sqrt{d}}\right)V, \quad Q = \mathbf{h}^{(\ell+1)}, \quad K = g_{\text{glob}}, \quad V = g_{\text{glob}}. \quad (6.3b)$$

where d is the attention head dimension.

This integration enables each node to access chamber-scale geometric context and strengthens the encoder's capacity to capture long-range mechanical correlations. Finally, the updated node embeddings yield two representations required for subsequent modules, a local graph latent that preserves spatially resolved geometric information for decoding, and a global latent that summarises

chamber-level mechanics and is later fused with the temporal representation.

$$Z_{\text{global}} = \text{MeanPool}(Z_{\text{local}}), \quad (6.4)$$

6.2.2.3 Temporal RNN Encoder

The ground truth P-V-t relation in Fig. 6.1 shows that, for any prescribed cavity volume, the corresponding P-t waveform is smooth and periodic over a cardiac cycle. This behaviour reflects the underlying mechanics. Once the chamber volume is fixed, pressure evolution is governed by the time-varying active stress, passive tissue response, and contraction relaxation dynamics, all of which evolve smoothly in time. Motivated by this observation, the temporal RNN encoder (Fig. 6.4) is designed to represent the P-t behaviour at a given loading volume through a compact latent trajectory.

During training, each batch contains sequences associated with different target volumes while spanning all time points in the cycle. For each time step $t_k, k = 1, 2, \dots, t_0$, the input vector $\mathbf{T}_k''$ encodes the prescribed chamber volume V_{target} , its deviation from the reference volume used in the Graph Fusion Encoder $\Delta V = V_{\text{target}} - V_{LV}$, and a Fourier representation of the cardiac phase $(\sin(2\pi t_k/t_0), \cos(2\pi t_k/t_0))$.

An initial MLP maps this input to a time-conditioned latent vector:

$$\mathbf{T}_k' = \text{MLP}_{\text{init}}(\mathbf{T}_k''), \quad k = 1, \dots, t_0. \quad (6.5)$$

Temporal evolution is modelled using a GRU, which provides a lightweight recurrent operator that captures the smooth and history-dependent nature of cardiac pressure dynamics. The first hidden state is initialised as:

$$\mathbf{T}_1 = \text{MLP}_{T_1}(\mathbf{T}_1'). \quad (6.6)$$

The subsequent states follow the standard GRU update rule:

$$\mathbf{T}_{k+1} = \text{GRU}(\mathbf{T}_{k+1}', \mathbf{T}_k), \quad k = 1, \dots, t_0 - 1. \quad (6.7)$$

Unrolling this recurrence over the full cardiac cycle yields the temporal latent representation at the final time step.

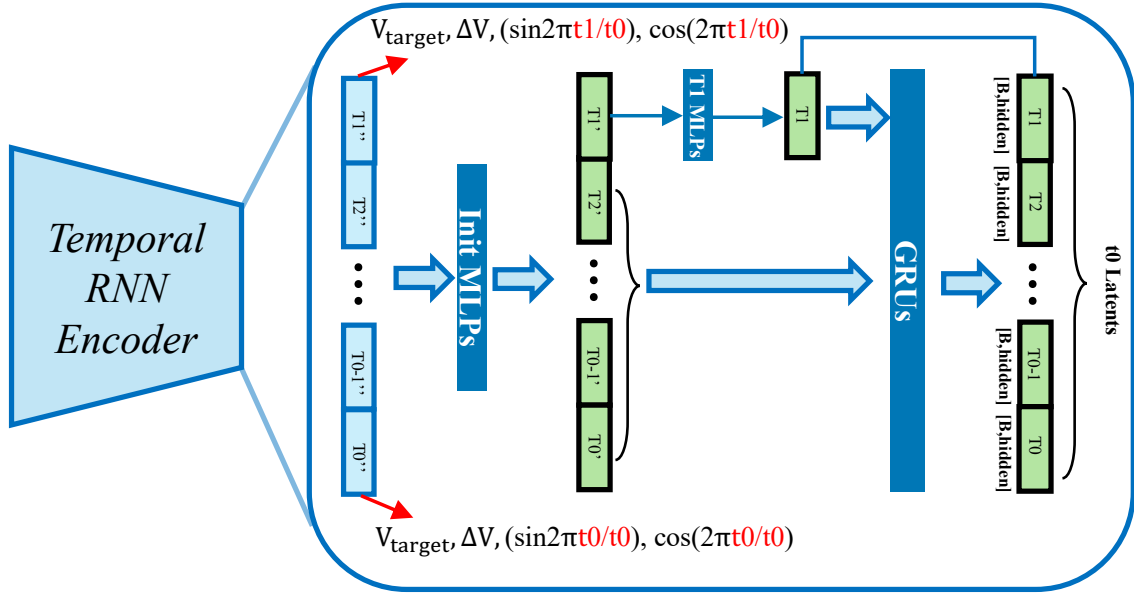


Figure 6.4: **Schematic of the Temporal RNN Encoder:** The input sequence encodes the target chamber volume, its deviation from the reference volume, and a Fourier time representation. An initial MLP generates time-conditioned latent vectors, which are propagated through a GRU to produce the final temporal latent used for latent fusion.

6.2.2.4 Latent Fusion Mechanism

As illustrated in Fig. 6.2, the latent representations produced by the Graph Fusion Encoder and the Temporal RNN Encoder are combined to serve as inputs to the displacement and pressure decoders. These two decoders require distinct types of information. The displacement decoder operates at the nodal level and therefore receives a local fusion latent $\mathbf{L}_{\text{local}} \in \mathbb{R}^{B \times N \times t_0 \times \text{hidden}}$, while the pressure decoder predicts a chamber level scalar quantity and uses a global fusion latent $\mathbf{L}_{\text{global}} \in \mathbb{R}^{B \times t_0 \times \text{hidden}}$.

To construct these representations, we first fuse the global graph latent $\mathbf{Z}_{\text{global}}$ with the temporal latent sequence $\{\mathbf{T}_k\}_{k=1}^{t_0}$. For each time step, the temporal vector $\mathbf{T}_k$ is added to the shared global descriptor and projected through a linear transformation, yielding:

$$\mathbf{L}_{\text{global}} = \{W_{\text{global}}(\mathbf{Z}_{\text{global}} + \mathbf{T}_k)\}_{k=1}^{t_0} \in \mathbb{R}^{B \times t_0 \times \text{hidden}} \quad (6.8)$$

This operation incorporates both the chamber-scale geometric information encoded in $\mathbf{Z}_{\text{global}}$ and the time-dependent loading behaviour captured by $\mathbf{T}_k$, providing the pressure decoder with a representation aligned with the global mechanics governing pressure evolution.

The local fusion latent $\mathbf{L}_{\text{local}}$ is constructed by combining the node-wise local graph latent

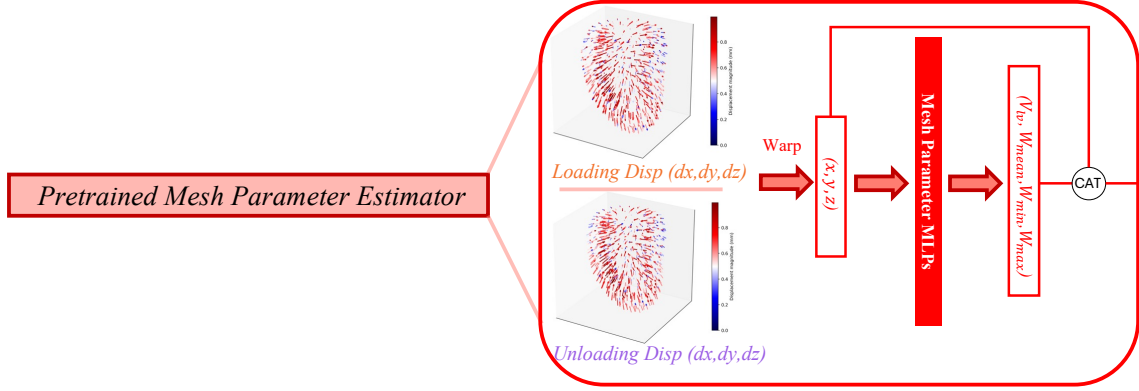


Figure 6.5: Pretrained MLP-based mesh parameter estimator used to compute global LV descriptors.

$Z_{\text{local}} \in \mathbb{R}^{B \times N \times \text{hidden}}$ with the global fusion latent $\mathbf{L}_{\text{global}} \in \mathbb{R}^{B \times t_0 \times \text{hidden}}$. Both tensors are first broadcast to the shared shape $B \times N \times t_0 \times \text{hidden}$, after which they are added element-wise and projected through a linear transformation:

$$\mathbf{L}_{\text{local}} = W_{\text{local}}(\mathfrak{B}(Z_{\text{local}}) + \mathfrak{B}(\mathbf{L}_{\text{global}})) \in \mathbb{R}^{B \times N \times t_0 \times \text{hidden}}. \quad (6.9)$$

This formulation injects temporal and chamber level loading information into every node while retaining spatially resolved geometric structure, providing a unified representation for both displacement and pressure prediction. In this way, the fused latent reflects the FEA principle that nodal deformation fields and chamber level pressure both depend on the interaction between local geometric properties and global loading constraints.

6.2.2.5 Cycle Consistency Strategy

Each training epoch performs three forward-inverse passes through the model to enforce a load-ing-unloading-loading cycle. This cycle consistency constrains the loading and unloading decoders to remain mutually coherent and operate within a mechanically consistent deformation space.

(1) Forward loading pass: The unloaded mesh $\mathbf{X}_{\text{unload}}^*$ and temporal descriptors are first processed by the encoder and loading decoders to produce the forward predictions $(P_t^{(1)}, \mathbf{u}_t)$, yielding the predicted loaded mesh $\mathbf{X}_{\text{load}}^{(1)}$. A pretrained MLP-based mesh parameter estimator then extracts the global descriptors $(V_{LV}, W_{\text{mean}}, W_{\text{min}}, W_{\text{max}})$ from this predicted loaded geometry (Fig. 6.5).

(2) Unloading pass: The predicted loaded mesh $\mathbf{X}_{\text{load}}^{(1)}$ is combined with the same temporal

descriptors by setting $V_{\text{target}} = V_{LV}$, so that the unloading decoder is queried at the volume of the predicted loaded state. This choice implies $\Delta V = 0$, which has no mechanical role during unloading but keeps the input format identical to that of the loading branch. The unloading decoder then produces $\mathbf{X}_{\text{unload}}^{(1)}$.

(3) Re-loading pass: The predicted unloaded mesh $\mathbf{X}_{\text{unload}}^{(1)}$ is passed again through the loading decoders to generate a second loading prediction $(P_t^{(2)}, \mathbf{X}_{\text{load}}^{(2)})$. Cycle consistency is enforced by comparing $\mathbf{X}_{\text{load}}^{(2)}$ with the forward prediction $\mathbf{X}_{\text{load}}^{(1)}$.

The total loss is:

$$\mathcal{L}_{\text{total}} = \lambda_p \mathcal{L}_p + \lambda_d \mathcal{L}_{\text{disp}}. \quad (6.10)$$

Both the pressure loss $\mathcal{L}_p$ and the displacement loss $\mathcal{L}_{\text{disp}}$ contain supervised observation and cycle consistency terms:

$$\mathcal{L}_p = \mathcal{L}_p^{\text{obs}} + \mathcal{L}_p^{\text{cycle}} = \|P_t^{(1)} - P^*\|_2^2 + \|P_t^{(2)} - P_t^{(1)}\|_2^2, \quad (6.11a)$$

$$\mathcal{L}_{\text{disp}} = \mathcal{L}_{\text{disp}}^{\text{obs}} + \mathcal{L}_{\text{disp}}^{\text{cycle}}, \quad (6.11b)$$

$$\mathcal{L}_{\text{disp}}^{\text{obs}} = \|\mathbf{X}_{\text{load}}^{(1)} - \mathbf{X}_{\text{load}}^*\|_2^2 + \|\mathbf{X}_{\text{unload}}^{(1)} - \mathbf{X}_{\text{unload}}^*\|_2^2, \quad (6.11c)$$

$$\mathcal{L}_{\text{disp}}^{\text{cycle}} = \|\mathbf{X}_{\text{load}}^{(2)} - \mathbf{X}_{\text{load}}^{(1)}\|_2^2. \quad (6.11d)$$

where λ_p and λ_d denote the loss weights, and the supervised terms are computed using the ground-truth meshes and pressure $(\mathbf{X}_{\text{load}}^*, \mathbf{X}_{\text{unload}}^*, P^*)$.

Together, these terms ensure that the model learns a physically coherent and invertible mapping between the unloaded and loaded states. In addition, this cycle consistency formulation has been shown in our experiments to substantially reduce the amount of required supervision while improving both geometric and pressure prediction accuracy.

6.2.2.6 Inference procedure

At inference time, *CGFENet* operates in two modes: (i) forward loading prediction from an unloaded geometry and (ii) inverse unloading estimation from a loaded geometry. The procedure is summarised in Algorithm 2.

Algorithm 2: Inference procedure of CardioGraphFENet

Mode A: Forward loading prediction**Input:** Unloaded mesh $\mathbf{X}_{\text{unload}}^*$, one or multiple target volumes V_{target} .**Output:** Predicted pressure and deformation trajectories over the cardiac cycle at the specified V_{target} .**1. Geometric preprocessing:**

- Assign endocardial, epicardial and interior labels to $\mathbf{X}_{\text{unload}}^*$.
- Extract V_{LV} , W_{mean} , W_{min} , W_{max} using the mesh parameter estimator and form the geometric node features.

2. Temporal descriptor construction:

- Compute $\Delta V = V_{\text{target}} - V_{\text{LV}}$.
- Construct $(V_{\text{target}}, \Delta V, \sin(2\pi t_k/t_0), \cos(2\pi t_k/t_0))$ using the fixed time samples (t_k/t_0) defined in training protocol.

3. Loading prediction:

- Apply the loading decoder of *CGFENet* to obtain $\{P_t(V_{\text{target}}, t_k), \mathbf{u}_t(V_{\text{target}}, t_k)\}_{k=1}^{t_0}$.
- Form loaded meshes $\mathbf{X}_{\text{load}}(t_k) = \mathbf{X}_{\text{unload}}^* + \mathbf{u}_t(V_{\text{target}}, t_k)$.

4. Volume sweep and physiological PV loop (optional):

- Specify a target-volume range and, for each V_{target} in this range, repeat Steps 2–3 to generate a dense P-V-t mapping.
 - Use the predicted pressure traces to drive a lumped-parameter Windkessel model and obtain a physiologically consistent P-V loop.
-

Mode B: Inverse unloading estimation**Input:** Loaded mesh $\mathbf{X}_{\text{load}}^*(t_{\text{given}})$ at a known cardiac phase t_{given} **Output:** Estimated unloaded mesh $\mathbf{X}_{\text{unload}}$ **1. Geometric preprocessing:**

- Assign endocardial, epicardial and interior labels to $\mathbf{X}_{\text{load}}^*$.
- Extract V_{LV} , W_{mean} , W_{min} , W_{max} using the mesh parameter estimator and form the geometric node features.

2. Temporal descriptor construction:

- No additional user input is required, as the normalised time samples (t_k/t_0) are fixed and the volume is automatically set to V_{LV} .

3. Unloading prediction:

- Apply the unloading decoder of *CGFENet* to obtain $\{\mathbf{X}_{\text{unload}}(t_k)\}_{k=1}^{t_0}$.

4. Selecting the zero-pressure configuration:

- Choose the frame corresponding to the desired cardiac phase t_{given} . For example, for ED, the Fourier embedding is (0, 1), which corresponds to t_0 in our temporal parametrisation. $\mathbf{X}_{\text{unload}}(t_0)$ is returned as the estimated zero-pressure state.
-

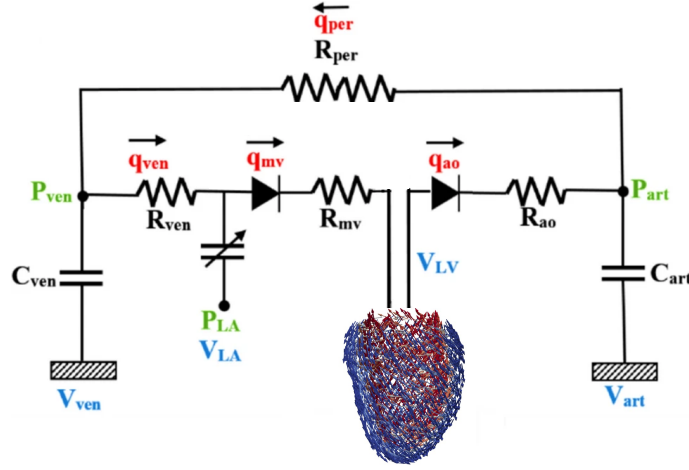


Figure 6.6: **Schematic of the lumped parameter model employed in the simulations.** C : capacitance (compliance); R : resistance.

6

6.2.3 Lumped parameter model coupling and P-V-t lookup table

A conventional approach to simulate the cardiac cycle involves coupling the FEA governing equations with a lumped parameter circulation model, typically a Windkessel model (Fig. 6.6). This hydromechanical analogue relates ventricular pressure to the ejection flow rate through arterial compliance and peripheral resistance. In such a coupled framework, cavity pressure and volume are not independent inputs. Instead, they are interdependent variables determined by the instantaneous equilibrium between the ventricle time-varying elastance and the vascular impedance. Therefore, finding the equilibrium state at each time step requires a computationally expensive nested iterative scheme. An outer loop updates pressure to satisfy global flow conservation, while an inner Newton Raphson loop solves for the corresponding equilibrium displacements.

To reduce coupling cost and enable use of the pre-trained model, we replace the conventional fully coupled scheme with a decoupled workflow that separates the computationally expensive mechanics solver from the Windkessel model. Specifically, the framework operates in two phases:

(1) Dense P-V-t table generation: In this phase, we use the pre-trained *CGFENet* to construct a dense lookup table that maps prescribed cavity volume and cardiac time to the corresponding chamber pressure and displacement field. For each subject specific unloaded mesh Ω_0 , we sample a physiologically plausible grid $\{(V_i, t_k)\}$ spanning the desired volume range from 40 mL to 160 mL and all time points in the cardiac cycle. Each query pair (V_i, t_k) is then evaluated by the *CGFENet* to obtain:

$$\{P_{i,k}, \mathbf{u}_{i,k}\} = \mathcal{G}_\theta(V_i, t_k, \Omega_0),$$

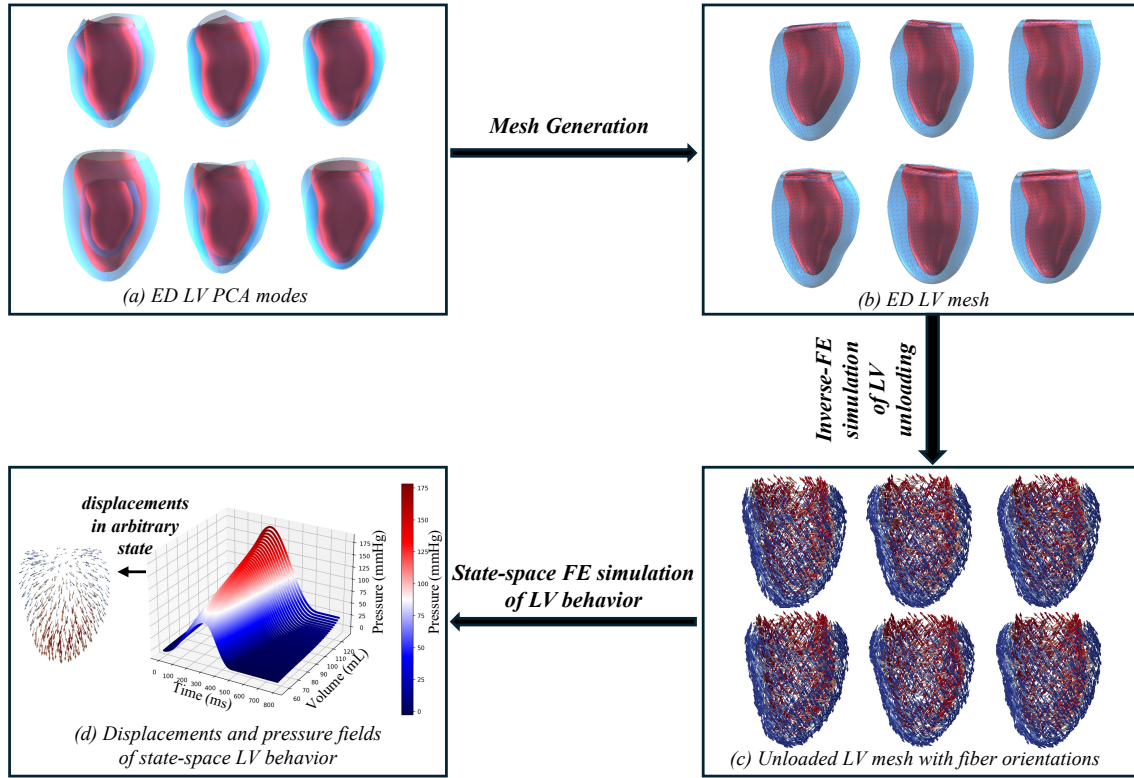


Figure 6.7: **Overview of the data-generation pipeline:** (a) PCA-based sampling of ED LV shapes, (b) construction of subject-specific tetrahedral ED meshes, (c) inverse-FEA computation yielding the unloaded, zero-pressure configuration with assigned fibre architecture, and (d) state-space FEA simulations producing displacement and pressure fields across prescribed cavity volumes.

where $\mathcal{G}_\theta$ denotes the learned mapping. Collecting all queries yields a dense P-V-t table that is independent of the circulation model and can be reused across different haemodynamic settings.

(2) Post hoc haemodynamic coupling: With the P-V-t table available, the lumped parameter circulation coupling is applied as a fast post-processing step. Instead of repeatedly calling a mechanics solver at each time step, the circulation update queries the P-V-t lookup table to obtain the pressure corresponding to the current volume and time. The circuit then updates the ejection flow and the next volume state, and the process is repeated over multiple cardiac cycles until the solution becomes periodic. This decoupled workflow removes the nested mechanics circulation iterations of a fully coupled scheme, enabling near-real-time reconstruction of the closed P-V loop.

6.2.4 Data Generation

The full data generation workflow is illustrated in Fig. 6.7, which includes (a) sampling ED LV geometries from a PCA shape model, (b) constructing volumetric ED meshes, (c) obtaining the cor-

responding unloaded configurations via inverse FEA simulation, and (d) generating displacement and pressure fields through state space FEA simulation.

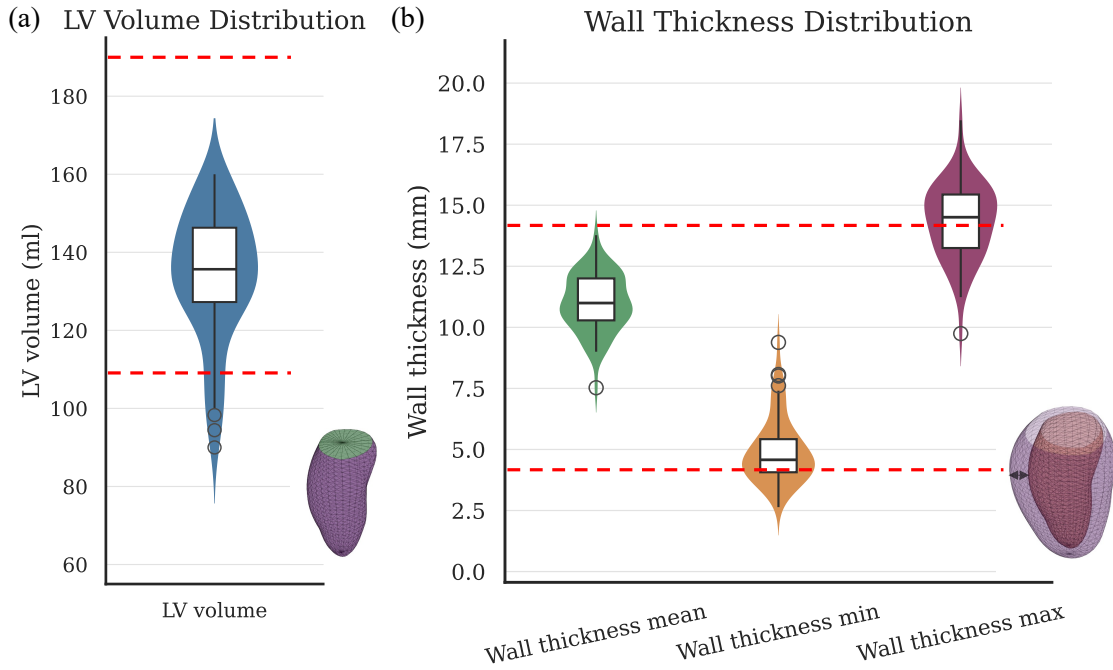


Figure 6.8: (a) **Distribution of ED LV cavity volumes** and (b) **myocardial wall thickness metrics (mean, minimum, maximum)** across all synthesized meshes. The three wall-thickness distributions represent statistics computed over the myocardial wall surface of the same heart. These distributions are included to demonstrate the physiological plausibility and reliability of the generated dataset, with the wall-thickness values remaining within ranges reported in prior studies. The red dashed lines denote reference ranges derived from the literature, including normal adult ED LV volumes (approximately 110–190 mL) [164] and ED myocardial wall thickness spanning roughly 3–13 mm [165].

(1) ED shape sampling and volumetric meshing: We first sampled ED LV geometries from the publicly available PCA statistical shape model provided by the *Cardiac Atlas Project (CAP)* (<https://www.cardiacatlas.org/left-ventricular-modes/>). The model was trained on 1,991 asymptomatic subjects and captures dominant ED anatomical variations, including global size, sphericity, and concentricity [154]. To ensure that the synthesised geometries remained within physiologically plausible limits, each PCA coefficient was truncated to the interval $[-2, 2]$, corresponding to approximately 95% of population level variability. For each sampled ED surface, a tetrahedral FEA mesh was generated using Gmsh. Owing to meshing variability, the resulting volumetric meshes exhibited heterogeneous resolutions, with node counts ranging from 400 to 1000 and tetrahedral elements from 1000 to 2000 (Fig. 6.7b). All meshes were visually inspected to ensure anatomical plausibility and sufficient geometric quality for nonlinear FEA simulation. The distributions of ED cavity volume and myocardial wall thickness across all synthesised meshes are

presented in Fig. 6.8, showing that the vast majority of shapes lie within physiologically reported ranges for healthy adults. A small number of samples fall marginally outside these intervals, which we retain to broaden geometric coverage and to provide controlled extreme cases that enhance the robustness of model training.

(2) Inverse FEA unloading: For each ED LV mesh, fibre orientations were first assigned, after which an inverse FEA unloading computation was performed to obtain its corresponding zero-pressure configuration. This step preserves a one to one anatomical correspondence between loaded and unloaded geometries (Fig. 6.7c).

(3) State space FEA simulation: Starting from each unloaded configuration, a state space FEA model was used to simulate LV mechanical behaviour across a range of cavity volumes. The resulting displacement and pressure fields (Fig. 6.7d) form a dense set of P-V-t states (lookup table) for supervising the loading and unloading branches of *CGFENet*.

A total of 67 anatomically distinct LV meshes were generated, yielding 1,814,146 P-V-t states. To avoid geometric leakage, the dataset was partitioned at the mesh level using an approximately 1:4 ratio for testing and training, with all states from each mesh retained within the same subset.

To avoid geometric leakage, where different P-V-t states generated from the same underlying LV geometry appear in both the training and testing sets and potentially lead to overly optimistic evaluation of geometric generalisation, the dataset was partitioned at the mesh level using an approximately 1:4 ratio for testing and training, with all states from each mesh retained within the same subset.

6.2.5 Experimental Setup

This section lists the network configurations (Table 6.1), FEA configurations (Table 6.2) and the evaluation metrics 6.3 used in the following experiments.

6.3 Results

6.3.1 Overview

The results section conducted a comprehensive experimental evaluation of *CGFENet*, focusing on three key aspects: accuracy, robustness, and physiological fidelity. The evaluation consisted of four

Table 6.1: **Key training configurations.**

Item	Setting
Hidden Channels	128
GATv2 Heads	4
Dropout Rate	0.1
Optimiser	AdamW
Learning Rate	1×10^{-3}
Weight Decay	1×10^{-3}
Loss weights	$\lambda_p = 1, \lambda_d = 10^4$
GPU	NVIDIA RTX 3090 (24 GB)

Table 6.2: **Key FE constitutive parameters.** The parameter settings are adapted from the literature [70] with minor adjustments.

Category	Value
Stiffness coefficient (C)	0.1 kPa
Material parameter (b_{ff})	29.9
Material parameter (b_{xx})	13.3
Material parameter (b_{fx})	26.6
Fiber angles (endo / epi)	$60^\circ / -60^\circ$
Peak active tension ($T_{\max}$)	67 kPa
Relaxed sarcomere length (l_r)	$1.85 \mu\text{m}$
Sarcomere length at which no active tension develops (l_0)	$1.58 \mu\text{m}$
Gradient of linear relaxation duration with sarcomere length relation (m)	$524 \text{ ms}\mu\text{m}$
Time intercept of linear relaxation duration with sarcomere length (b)	-750 ms
Time to peak tension (t_0)	300.5 ms
Cardiac Cycle	800 ms

parts. First, we benchmarked our method against several representative GNN baselines. Second, we conducted ablation studies to quantify the effect of each major component. Third, we investigated the role of cycle consistency in enabling reduced supervision. Finally, to assess physiological plausibility, we fed the predicted ventricular pressures into a lumped parameter Windkessel model and compared the resulting P-V loops against reference data from FEA simulations.

6.3.2 Baseline Comparison

We first benchmarked *CGFENet* against two representative graph based surrogate models, GraphUNet [139] and MeshGraphNet [137], for the forward loading prediction task. As reported in Table 6.4, *CGFENet* achieved the best overall geometric accuracy, yielding the lowest displacement Root Mean Squared Error (RMSE), the highest vertex level accuracy, and a substantially reduced HD. In particular, our method lowered the HD to $1.89 \pm 0.15 \text{ mm}$, improving vertex level fidelity by around 50% relative to the baselines.

For pressure prediction, our model also achieved the strongest agreement with the FEA refer-

Table 6.3: Evaluation metrics used for geometric and pressure-based assessment.

Metric	Description
Vertex overlap accuracy	Percentage of mesh nodes with Euclidean error < 1 mm or 2 mm. This metric assesses fine-scale geometric agreement and is sensitive to local surface discrepancies under fixed node correspondence.
Root Mean Square Error (RMSE)	Root mean squared Euclidean distance between corresponding nodes. It quantifies the average geometric reconstruction error and reflects overall mesh fidelity.
Hausdorff distance (HD)	Maximum Euclidean deviation between two surfaces. It measures worst-case geometric deviation and is sensitive to localised outliers and boundary discrepancies.
Coefficient of determination (R^2)	Statistical measure of agreement between predicted and reference pressure values. It evaluates the agreement between predicted and reference pressure profiles.

ences, attaining the highest R^2 (0.98) and the lowest RMSE (1.70 mmHg). The P-t trajectories at several fixed target cavity volumes (Fig. 6.9) further highlighted these differences, where each continuous curve was generated under the same prescribed target volume. The ground truth curves exhibited smooth and physiologically plausible systolic upstrokes followed by diastolic relaxation, and these temporal patterns were closely preserved by *CGFENet*, which tracked both timing and peak pressure with consistently small errors across the cycle. In contrast, MeshGraphNet produced relatively smooth curves but showed clear bias during the rapid rising pressure phase and around peak systole. GraphUNet performed worst, exhibiting pronounced oscillations and non smooth fluctuations throughout the cycle, which indicated poor temporal coherence across frames. Overall, explicitly incorporating temporal dynamics enabled *CGFENet* to generate stable and physiologically consistent pressure trajectories under complex deformation states, whereas the baselines either deviated during fast transients (MeshGraphNet) or became temporally unstable (GraphUNet).

In terms of runtime and model size, we compared *CGFENet* with MeshGraphNet and conventional FEA (Table 6.5). *CGFENet* used substantially fewer learnable parameters, with more than a two fold reduction, because the temporal RNN module was decoupled from the graph encoder. This avoided a large increase in graph processing capacity while still modelling time dependence. This lightweight design also translated into practical efficiency. *CGFENet* generated the full candidate P-V-t map for a case in under one minute, which was approximately $80\times$ faster than the traditional FEA pipeline. This acceleration reduced the turnaround time of cycle wide simulations and represents an important step toward clinically feasible deployment.

Overall, these results demonstrate that *CGFENet* serves as a more accurate surrogate for

Table 6.4: Comparison with representative graph-based surrogate models on forward loading prediction. FEA served as ground truth.

Model	Displacement			Pressure	
	RMSE (mm)↓	Vertex Acc. < 1 mm ↑	HD (mm)↓	R^2 ↑	RMSE (mmHg)↓
GraphUNet	0.42 ± 0.02	0.85 ± 0.01	3.30 ± 0.11	0.95 ± 0.01	5.83 ± 0.22
MeshGraphNet	0.44 ± 0.02	0.81 ± 0.03	4.88 ± 0.24	0.98 ± 0.01	3.63 ± 0.10
<i>CGFENet</i> (ours)	0.39 ± 0.01	0.90 ± 0.01	1.89 ± 0.15	0.98 ± 0.01	1.70 ± 0.17

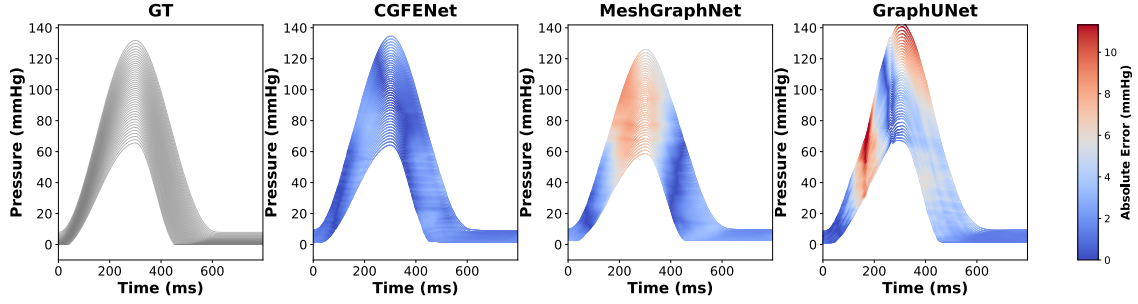


Figure 6.9: P-t trajectories for a representative test case under several prescribed target volumes, comparing the ground-truth FEA solution (GT) with predictions from *CGFENet*, MeshGraphNet, and GraphUNet. In each panel, *CGFENet* most closely follows the target curve and yields the smallest volume errors, while the baseline surrogates exhibit larger deviations from the ground truth.

myocardial loading mechanics, while maintaining stronger temporal consistency and more physiologically coherent pressure dynamics than the compared graph based architectures.

6.3.3 Ablation Study

We performed a comprehensive ablation study to quantify the contribution of each major design choice in *CGFENet*. Tables 6.6 and 6.7 report the results for (i) displacement prediction in both the forward loading and inverse unloading directions, and (ii) pressure prediction, respectively. Overall, the full model (V0) provided the best balanced performance across geometric and haemodynamic metrics, while most ablations degraded either accuracy, vertex fidelity, or pressure error.

(1) Temporal modelling: Replacing the recurrent temporal module with an MLP based temporal encoder (V1) produced mixed outcomes. In terms of displacement, V1 marginally reduced the average RMSE in both loading and unloading, but it markedly worsened vertex level fidelity, with the HD increasing, especially for loading (3.13 mm vs. 1.89 mm for V0). For pressure prediction, the same modification consistently increased RMSE (2.43 vs. 1.70 mmHg). This pattern suggests that an MLP can match mean vertex wise errors but is less effective at suppressing local geometric outliers and capturing rapid transient dynamics, whereas recurrent temporal modelling

Table 6.5: **Comparison of model size and inference runtime for generating a full case-level P-V-T map.**

Model	Total Parameters ↓	Time (s) ↓
CardioGraphFENet (ours)	0.97 M	44.32 ± 11.23
MeshGraphNet	2.29 M	60.58 ± 13.83
FEM	—	3256.82 ± 1599.17

better preserves the systolic upstroke and peak pressure behaviour.

(2) Global fusion: Removing the global fusion module (V2) consistently degraded performance in both directions. Displacement errors increased and vertex level fidelity deteriorated substantially, with the HD rising to 3.31 mm for loading and 1.75 mm for unloading (vs. 1.89 mm and 1.38 mm in V0). Pressure accuracy also dropped, with RMSE increasing to 2.12 mmHg. Overall, these results indicate that injecting global contextual information is important for reproducing FEA like behaviour, supporting the design rationale described in Section 2.2.2.

(3) Cycle consistency constraint: Disabling the cycle consistency constraint (V3) exhibited a clear direction dependent impact. For forward loading, V3 achieved comparable average displacement errors and even a slightly lower HD. However, for inverse unloading it substantially degraded performance, with RMSE increasing from 0.30 to 0.41 mm and HD more than doubling (1.38 to 3.21 mm). Pressure prediction also worsened, with RMSE rising from 1.70 to 2.03 mmHg. Overall, these results suggest that cycle consistency stabilises the inverse mapping and mitigates accumulated drift, thereby improving unloading accuracy and maintaining better pressure fidelity.

(4) Graph backbone: Replacing the GATv2 backbone with alternative message passing operators consistently reduced performance. Switching to a GCN (V4) led to one of the largest drops in pressure accuracy (RMSE 3.11 mmHg) and markedly worsened forward displacement fidelity (HD 3.42 mm). Using GAT (V5) [149] and GraphSAGE (V6) yielded intermediate results, but both remained inferior to V0 across most metrics. Overall, these trends suggest that the stronger attention mechanism in GATv2 better captures heterogeneous mechanical interactions over the tetrahedral mesh, which benefits both displacement prediction and pressure estimation.

(5) Pre-trained mesh estimator: We further evaluated the effect of the mesh estimator that provided reference volume and wall thickness descriptors. In V7, the estimator was trained jointly with the full network, instead of being pre-trained using the unloaded dataset, which reduced performance across both geometry and pressure metrics, for example pressure RMSE increased from 1.70 to 3.24 mmHg. In V8, removing the mesh estimator entirely, thereby discarding the reference volume and wall thickness inputs, led to the most pronounced degradation, especially in

forward loading, where vertex level fidelity deteriorated markedly, with HD increasing to 4.88 mm, and pressure RMSE further rising to 3.75 mmHg. These results indicate that the LV volume and wall thickness descriptors provide effective global conditioning signals that improve both geometric fidelity and pressure prediction, and that pre-training the estimator yields more reliable cues than learning it purely end to end.

In summary, the ablation results demonstrate that *CGFENet* accuracy and stability stem from the combined contributions of explicit temporal modelling, global fusion, cycle consistency regularisation, an attention based graph backbone, and a pre-trained mesh estimator. Although some variants matched or slightly improved a single averaged metric, these gains typically came at the expense of poorer vertex level fidelity and or reduced pressure accuracy, underscoring the importance of the complete design for balanced performance.

Table 6.6: **Displacement Ablation**

Forward (Loading)					
Model	Description	RMSE (mm)↓	Vertex Acc. < 1mm ↑	Vertex Acc. < 2mm ↑	HD (mm)↓
V0	Full model (baseline)	0.39 ± 0.01	0.90 ± 0.01	0.99 ± 0.02	1.89 ± 0.15
V1	RNN → MLP temporal encoder	0.34 ± 0.01	0.89 ± 0.01	0.99 ± 0.02	3.13 ± 0.19
V2	w/o global fusion module	0.43 ± 0.01	0.84 ± 0.02	0.98 ± 0.02	3.31 ± 0.22
V3	w/o cycle-consistency constraint	0.38 ± 0.01	0.88 ± 0.01	0.98 ± 0.01	1.58 ± 0.12
V4	GATv2 → GCN backbone	0.44 ± 0.01	0.84 ± 0.01	0.98 ± 0.02	3.42 ± 0.19
V5	GATv2 → GAT backbone	0.42 ± 0.01	0.85 ± 0.01	0.99 ± 0.02	3.06 ± 0.18
V6	GATv2 → GraphSAGE backbone	0.47 ± 0.01	0.86 ± 0.02	0.98 ± 0.02	3.57 ± 0.16
V7	w/o pretrained mesh estimator	0.40 ± 0.01	0.87 ± 0.01	0.99 ± 0.02	3.02 ± 0.20
V8	w/o mesh estimator	0.47 ± 0.02	0.86 ± 0.01	0.97 ± 0.04	4.88 ± 0.24
Inverse (Unloading)					
Model	Description	RMSE (mm)↓	Vertex Acc. < 1mm ↑	Vertex Acc. < 2mm ↑	HD (mm)↓
V0	Full model (baseline)	0.30 ± 0.02	0.93 ± 0.01	0.99 ± 0.03	1.38 ± 0.12
V1	RNN → MLP temporal encoder	0.20 ± 0.02	0.92 ± 0.01	0.99 ± 0.03	1.52 ± 0.11
V2	w/o global fusion module	0.37 ± 0.02	0.86 ± 0.02	0.98 ± 0.06	1.75 ± 0.12
V3	w/o cycle-consistency constraint	0.41 ± 0.03	0.89 ± 0.02	0.97 ± 0.08	3.21 ± 0.12
V4	GATv2 → GCN backbone	0.29 ± 0.02	0.92 ± 0.01	0.99 ± 0.04	1.58 ± 0.12
V5	GATv2 → GAT backbone	0.32 ± 0.02	0.89 ± 0.02	0.98 ± 0.05	1.67 ± 0.12
V6	GATv2 → GraphSAGE backbone	0.30 ± 0.03	0.92 ± 0.01	0.99 ± 0.06	1.54 ± 0.11
V7	w/o pretrained mesh estimator	0.39 ± 0.04	0.84 ± 0.03	0.97 ± 0.03	1.76 ± 0.14
V8	w/o mesh estimator	0.29 ± 0.02	0.84 ± 0.01	0.98 ± 0.02	1.51 ± 0.10

Table 6.7: **Pressure Ablation**

Model	Description	$R^2\uparrow$	RMSE (mmHg) $\downarrow$
V0	Full model (baseline)	0.98 ± 0.01	1.70 ± 0.17
V1	RNN $\rightarrow$ MLP temporal encoder	0.98 ± 0.01	2.43 ± 0.18
V2	w/o global fusion module	0.98 ± 0.01	2.12 ± 0.18
V3	w/o cycle consistency constraint	0.98 ± 0.01	2.03 ± 0.21
V4	GATv2 $\rightarrow$ GCN back-bone	0.98 ± 0.01	3.11 ± 0.14
V5	GATv2 $\rightarrow$ GAT back-bone	0.98 ± 0.01	2.41 ± 0.16
V6	GATv2 $\rightarrow$ GraphSAGE backbone	0.97 ± 0.01	2.78 ± 0.14
V7	w/o pre-trained mesh estimator	0.98 ± 0.01	3.24 ± 0.15
V8	w/o mesh estimator	0.98 ± 0.01	3.75 ± 0.10

6.3.4 Reduced supervision capability

To verify the effectiveness of the cycle consistency design under limited labels, we conducted a reduced supervision study in which the SR controls the fraction of training cases with ground truth FEA labels. Importantly, all training cases, labelled and unlabelled, were still used during training. Labelled cases contributed both supervised losses and the cycle consistency loss, whereas unlabelled cases contributed only through the cycle consistency objective. We report the loading vertex level accuracy, defined as the fraction of vertices with displacement error below 1 mm, over training epochs in Fig. 6.10, comparing *Cycle* against *No cycle*.

With cycle consistency enabled, performance degraded only mildly as SR decreased. The SR = 70%, 50%, and 30% curves converged to similarly high plateaus of approximately 0.87 to 0.90, indicating that the cycle loss can exploit the full training set even when explicit labels are missing. Even at SR = 10%, the model maintained a stable plateau around 0.70, demonstrating substantial label efficiency.

In contrast, removing cycle consistency caused a pronounced collapse under low SR. While SR = 70% and 50% remained competitive, accuracy dropped sharply at SR = 30% and deteriorated further at SR = 10%, plateauing near 0.25. Overall, these results confirm that cycle consistency provides an effective reduced and self supervised learning signal, allowing *CGFENet* to leverage all training geometries and maintain vertex level fidelity under substantially reduced labelled supervision.

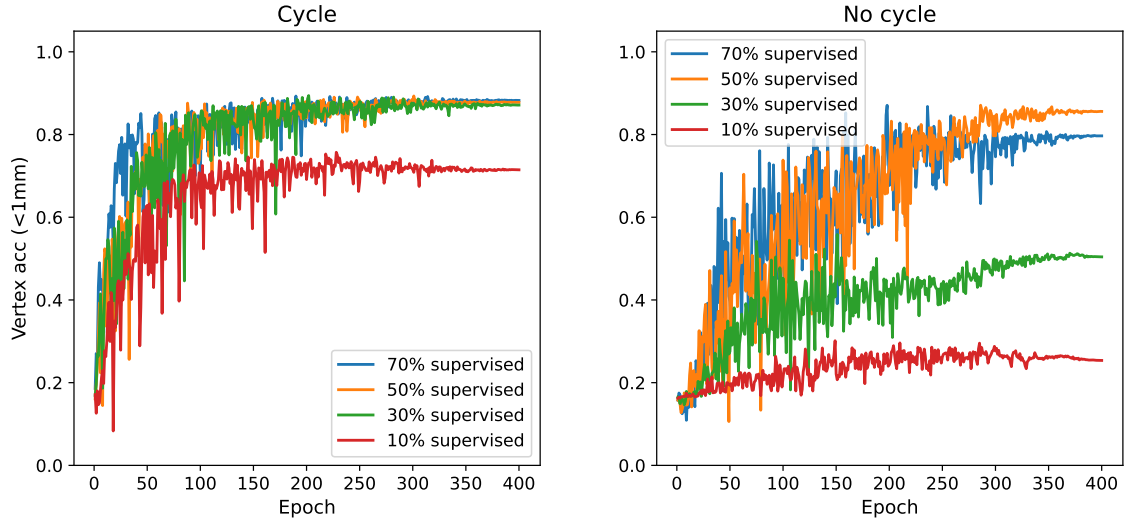


Figure 6.10: **Reduced-supervision study:** Vertex-level accuracy versus training epoch under SR of 70%, 50%, 30%, and 10%. *Cycle* uses all training cases: labelled samples contribute supervised and cycle-consistency losses, while unlabelled samples contribute only through cycle consistency. *No cycle* also processes all cases but updates the network using labelled samples only (unlabelled samples provide no loss/gradients).

6.3.5 Windkessel-coupled P-V loop evaluation

To provide a case level interpretation of model performance, we selected four representative test cases with distinct anatomy. Table 6.8 reports their unloaded LV cavity volume and wall thickness descriptors (V_{LV} , W_{mean} , W_{min} , W_{max}), spanning a substantial range of cavity sizes ($V_{LV} = 67.12\text{--}84.63\text{ mL}$) and thickness heterogeneity ($W_{\text{mean}} = 10.61\text{--}12.30\text{ mm}$, $W_{\text{min}} = 3.67\text{--}8.01\text{ mm}$, $W_{\text{max}} = 13.88\text{--}16.03\text{ mm}$).

We generated full-cycle P-V loops by coupling *CGFENet* with the Windkessel model introduced in Section 2.3.2. The surrogate was queried at candidate (V_t, t) states to predict pressure P_t and the associated deformation, while the Windkessel dynamics updated the cavity volume trajectory $V(t)$. This surrogate Windkessel iteration was repeated until the $(P(t), V(t))$ trajectory converged to a stable closed P-V loop.

Fig. 6.11 compares the converged P-V loops against FEA results for the four cases. The surrogate reproduced both systolic and diastolic branches closely, with only minor deviations in Case 3, mainly around transition regions. From these loops, we also computed EDV, ESV, and EF (Table 6.9). The absolute differences to FEA remained small across all cases, with $\Delta\text{EDV} = 0.07\text{--}1.11\text{ mL}$, $\Delta\text{ESV} = 0.01\text{--}1.82\text{ mL}$, and $\Delta\text{EF} = 0.01\text{--}0.93\text{ pp}$, and the largest discrepancies were again observed in Case 3.

Table 6.8: **Unloaded state LV cavity volume and myocardial wall-thickness descriptors** for four representative cases, including cavity volume V_{lv} and wall thickness statistics (W_{mean} , W_{min} , W_{max}) computed from endocardial–epicardial distances.

Case ID	V_{lv} (mL)	W_{mean} (mm)	W_{min} (mm)	W_{max} (mm)
Case 1	67.12	10.61	8.01	13.88
Case 2	70.84	12.30	7.61	16.03
Case 3	81.71	12.01	4.58	15.45
Case 4	84.63	11.18	3.67	15.60

Table 6.9: **Global LV functional indices derived from the Windkessel-coupled P-V loops** for four representative cases. EDV: end-diastolic volume, ESV: end-systolic volume, EF: ejection fraction. Δ denotes the absolute difference between *CGFENet* and FEA, with ΔEF reported in percentage points (pp).

Case	EDV (mL)	ΔEDV (mL)	ESV (mL)	ΔESV (mL)	EF (%)	ΔEF (pp)
Case 1	115.12	0.20	66.23	0.05	42.47	0.06
Case 2	121.22	0.08	61.07	0.03	49.62	0.01
Case 3	131.18	1.11	70.35	1.82	46.37	0.93
Case 4	145.09	0.07	84.11	0.01	42.03	0.02

Fig. 6.12 further provides qualitative geometry comparisons at ED and ES for the four representative cases. For each case and phase, the colour panels of *CGFENet* and FEA visualise the deformation magnitude with respect to the unloaded reference configuration, while the *Diff* panels show the prediction error relative to FEA (prediction minus FEA), thereby separating model residuals from the underlying physiological deformation. Across all cases, the predicted deformation patterns closely match the FEA references at both ED and ES, and the *Diff* maps are nearly uniformly blue, indicating consistently small discrepancies. Importantly, the deformation is not spatially uniform. The apex exhibits the largest motion as expected, and this regional pattern is consistently reproduced by the surrogate, supporting physiologically plausible mechanics rather than a trivial near rigid or globally uniform deformation.

6.4 Discussion

6.4.1 Overall performance and physiological fidelity

Overall, the results indicated that *CGFENet* provided a strong surrogate for LV loading and unloading mechanics, combining high geometric fidelity with accurate and temporally coherent pressure prediction. Compared with GraphUNet and MeshGraphNet, it achieved the best balanced performance, notably reducing vertex level error while maintaining strong pressure agreement with

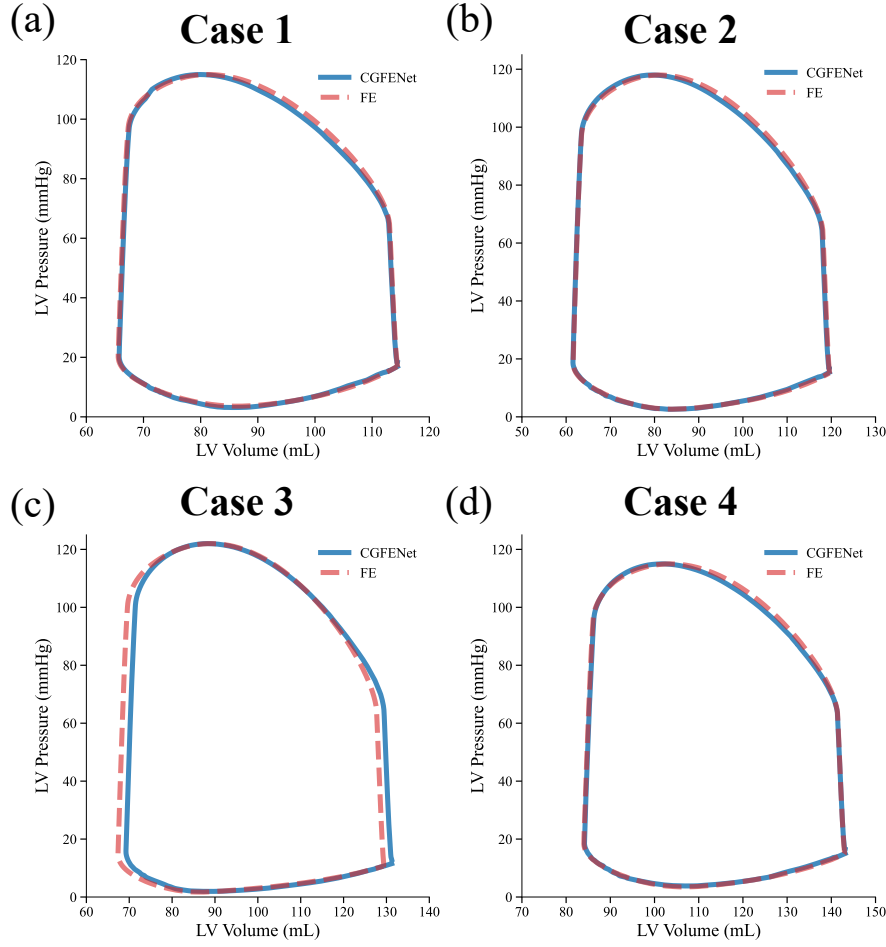


Figure 6.11: **Windkessel-coupled full-cycle P-V loops for four representative cases:** *CGFENet* predictions are compared against FEA references after iterative surrogate-Windkessel coupling until a stable closed loop is reached.

FEA and avoiding the oscillations or systolic phase bias observed in the baselines. The ablations further suggested that these gains were not attributable to a single component. Rather, they arose from the combined effects of temporal modelling, global fusion, and an attention based graph backbone, with cycle consistency being critical for stabilising the inverse unloading mapping and enabling label efficient training under reduced supervision. Furthermore, beyond pointwise metrics, the Windkessel coupled evaluation showed that the predicted pressures supported physiologically plausible closed P-V loops and that derived functional indices closely matched FEA across diverse cases, while maintaining expected regional deformation patterns. Finally, the model remained practically efficient, generating a full case level P-V-t map in under one minute, about $80\times$ faster than FEA, supporting its use in cycle wide simulation workflows where conventional FEA runtimes are prohibitive.

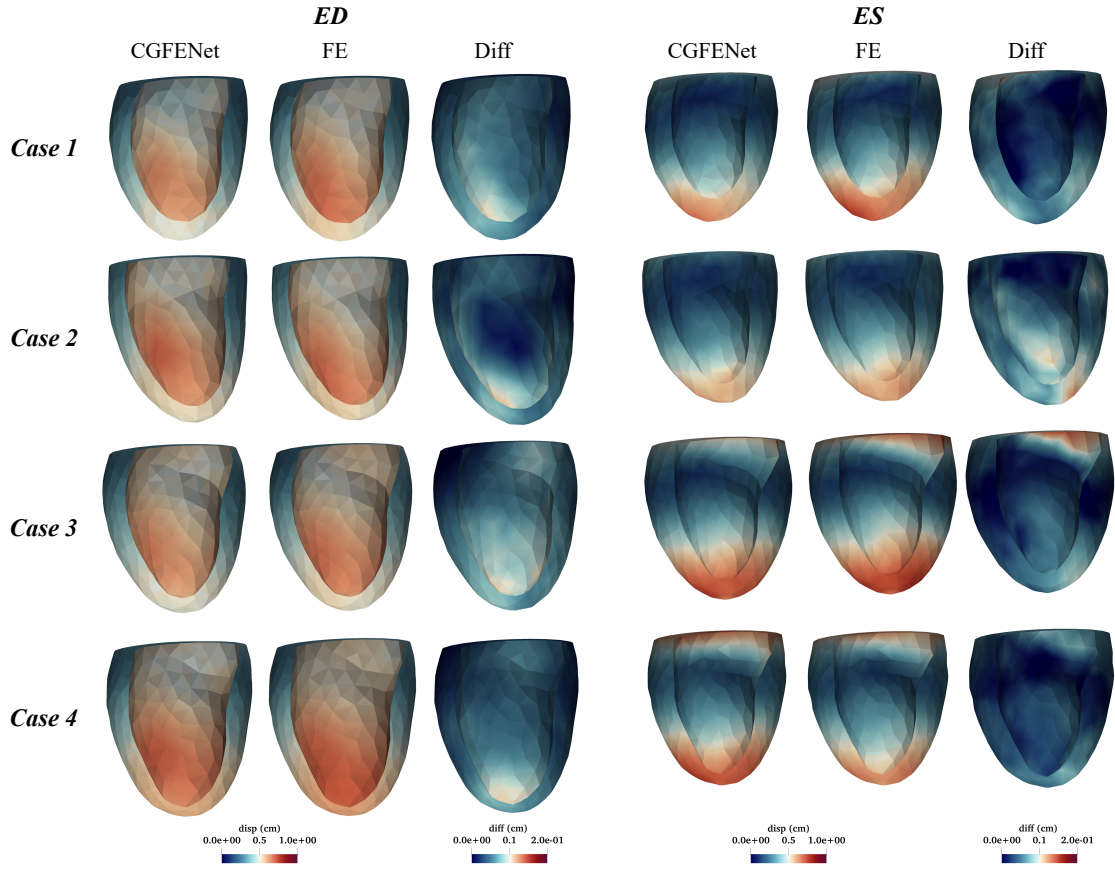


Figure 6.12: **Qualitative comparisons of ED and ES deformation fields** for four representative cases. Colour maps (*CGFENet* and FEA) show the deformation magnitude relative to the unloaded reference configuration, while *Diff* reports the residual error (prediction minus FEA).

6.4.2 Comparison with existing cardiac surrogate literature

Table 6.10 and 6.11 positions *CGFENet* relative to representative cardiac surrogate models in terms of dataset coverage and modelling formulation. Current graph based deep surrogates largely focus on passive only mechanics under pressure driven loading, either with fixed meshes, for example Dalton et al. [18], or with variable meshes, for example HeartSimSage [20], and typically learn a forward mapping $p \rightarrow \mathbf{u}$ and report performance on passive deformation states. By contrast, current full-cycle surrogates largely based on PINN style optimisation, for example Buoso et al. [22] and Motiwale et al. [21], incorporate a prescribed pressure and active tension signal and learn $(p, T_a) \rightarrow \mathbf{u}$, but are commonly trained in a patient specific manner and remain restricted to fixed topology meshes and loading only prediction, which limits scalability and reuse across anatomies. In comparison, *CGFENet* is a purely data driven graph surrogate, not a PINN, which enables efficient pre-training on a substantially larger and more diverse dataset of 1,814,146 full-cycle LV states with variable meshes. This variable mesh design is practically important. Using a GraphNet

style message passing formulation, the model operates on mesh connectivity without requiring node wise correspondence, allowing a single pre-trained network to generalise across different resolutions rather than relying on re-meshing to a fixed template. Moreover, rather than prescribing invasive pressure or relying on Windkessel coupling to generate pressure inputs, we adopt an imaging aligned volume driven formulation, $(V, t) \rightarrow (\mathbf{u}, p)$, where V can be readily extracted from medical images, for example segmentation derived cavity volumes. Overall, these choices extend prior patient specific, fixed mesh, loading only surrogates toward a reusable, cycle consistent framework that supports both loading and unloading and enables fast cycle wide inference across anatomies.

6.4.3 Limitations and future work

6

Despite the strong accuracy and physiological fidelity demonstrated in our experiments, this study has several limitations that motivate future extensions. First, the current formulation assumes a fixed set of myocardial parameter settings during data generation and training, such that only the state variables (V, t) are used to condition the predicted deformation and pressure. Therefore, the present surrogate does not explicitly model inter subject variability arising from changes in constitutive properties or activation, which may become important when translating to broader cohorts or pathological conditions. In future work, we will extend the framework to incorporate additional controllable parameter inputs, enabling direct parameter conditioned inference. Examples include myocardial stiffness parameters, such as global or regional stiffness scaling, maximal active tension or activation timing parameters, and fibre architecture descriptors such as helix angle distributions. Incorporating these variables would allow the surrogate to serve not only as a fast forward and inverse solver, but also as a differentiable engine for sensitivity analysis and parameter identification. Second, while the current results validate cycle level plausibility through Windkessel coupled P-V loops, further evaluation against more diverse anatomies and parameter regimes will be required to fully characterise robustness, particularly under extreme geometries or strongly altered material and activation settings. Addressing these limitations will broaden the clinical and research applicability of *CGFENet* toward rapid, controllable, and patient adaptive cardiac mechanics simulation.

Table 6.10: Comparison across representative cardiac surrogate models at the data level

Model	Dataset	Cycle	Mesh
Dalton et al. [18]	2,967 passive LV states	Passive only	Fixed
Shi et al. [20]	5,000 passive BiV states	Passive only	Variable
Motiwale et al. [21]	91 full-cycle LV states	Full-cycle	Fixed
Buoso et al. [22]	400 full-cycle LV states	Full-cycle	Fixed
Ours	1,814,146 full-cycle LV states	Full-cycle	Variable

Table 6.11: Comparison across representative cardiac surrogate models at the model level

Model	Direction	Mapping	Training
Dalton et al. [18]	Loading only	$p \rightarrow \mathbf{u}$	Pre-trained
Shi et al. [20]	Loading only	$p \rightarrow \mathbf{u}$	Pre-trained
Motiwale et al. [21]	Loading only	$(p, T_a) \rightarrow \mathbf{u}$	Patient specific
Buoso et al. [22]	Loading only	$(p, T_a) \rightarrow \mathbf{u}$	Patient specific
Ours	Loading and unloading	$(V, t) \rightarrow (\mathbf{u}, p)$	Pre-trained

6.5 Conclusion

We presented *CGFENet*, a unified DL surrogate for full-cycle LV mechanics that jointly supports forward loading and inverse unloading on variable tetrahedral meshes. By adopting an imaging aligned, volume driven formulation $(V, t) \rightarrow (\mathbf{u}, p)$ and combining attention based graph message passing with temporal modelling and cycle consistent training, the proposed framework achieves FEA level fidelity in both geometry and pressure while enabling efficient cycle wide inference. Across extensive evaluations, *CGFENet* outperforms representative graph based surrogates, maintains stable and physiologically coherent pressure trajectories, and generates complete case level P-V-t maps in under one minute, providing an approximately $80\times$ acceleration over conventional FEA. When coupled with a Windkessel model, the predicted pressures produce physiologically plausible closed P-V loops and accurate global functional indices across diverse anatomies. These results establish *CGFENet* as a practical and scalable surrogate for rapid cardiac mechanics simulation and downstream haemodynamic analysis.

Contents

7.1 Summary of contributions	126
7.2 Limitations and implications	127
7.3 Future directions	128
7.4 Closing remarks	128

This thesis aimed to advance image-driven LV biomechanics by improving two key stages of the personalised modelling pipeline: (i) robust 2D+t and 3D+t cardiac motion tracking across imaging modalities, and (ii) comprehensive DL surrogates that accelerate FEA while preserving full-cycle fidelity, including active contraction and inverse unloading capability. The thesis was structured around four aims spanning cross-modal registration, reduced-order physics-informed surrogates, unloaded-state estimation, and pre-trained full-cycle graph surrogates.

7.1 Summary of contributions

The main contributions of this thesis are summarised across four core methodological developments.

Unsupervised Cross-Modal Motion Tracking over the Cardiac Cycle (Chapter 3).

This thesis developed a unified registration framework that scales from 2D+t to 3D+t cardiac motion tracking and explicitly enforces cycle-level temporal consistency through a B-spline deformation parameterisation with DL Fourier-based regularisation. Across MRI and ECHO modalities, the framework produces robust full-cycle motion trajectories with practical computational runtimes, providing reliable motion inputs for downstream cardiac biomechanics modelling.

Reduced-Order Physics-Informed Neural Networks for Left Ventricular Mechanics over the Cardiac Cycle (Chapter 4). Building on the estimated cardiac motion, this thesis introduced an image-motion-consistent PINN–FEA formulation (*IMC-PINN-FE*) for patient-specific LV biomechanics. The framework jointly enforces biomechanical constraints and image-derived motion consistency, enabling data-efficient estimation of subject-specific myocardial stiffness and active tension while reconstructing physiologically coherent P-V dynamics. A key finding is that the proposed formulation identifies displacement solutions that remain physically valid while staying close to, rather than being strictly constrained by, potentially noisy image-derived motion. This improves robustness compared with fixed-tracking baselines.

Graph Neural Networks for Left Ventricular Unloading (Chapter 5). To address the long-standing computational bottleneck of inverse ventricular unloading, this thesis proposed a graph-based strategy to predict the zero-pressure reference configuration directly from diastolic geometries. This approach reduces reliance on iterative inverse FEA procedures and enables a more efficient end-to-end pipeline for patient-specific cardiac mechanics modelling.

Graph Neural Network Surrogates for Left Ventricular Mechanics over the Cardiac Cycle (Chapter 6). Finally, this thesis presented a unified pre-trained graph surrogate, *CGFENet*, that predicts full-cycle displacement and pressure fields from an imaging-aligned, volume-driven mapping $(V, t) \rightarrow (\mathbf{u}, p)$. Through attention-based graph message passing, temporal modelling, and cycle-consistent training, the framework supports both forward loading and inverse unloading on variable tetrahedral meshes. The model achieves FEA-level fidelity while enabling efficient construction of case-level P-V-t maps and substantial acceleration compared with conventional simulations. When coupled with a Windkessel model, the predicted pressures produce physiologically plausible closed P-V loops and accurate global functional indices.

7.2 Limitations and implications

Despite these advances, several limitations remain. First, 3D+t tracking and learning-based modelling are still constrained by dataset scale and modality-specific artefacts (e.g., speckle noise, dropouts, and variable acquisition protocols), which can reduce generalisation and temporal stability in difficult cases.

Second, while IMC-PINN-FE reproduces case-level P-V dynamics well, isovolumetric phases may be weakly expressed when the input V-t trajectory is smoothed by the imaging and tracking

pipeline. The thesis showed that these phases can be recovered when the prescribed V-t contains near-constant-volume segments (e.g. via Windkessel coupling), indicating that the limitation is primarily driven by the input volume signal rather than the PINN–FEA coupling itself. Furthermore, although uncertainty-aware temporal constraints were not explicitly formulated as dedicated loss terms, uncertainty estimation was partially incorporated through Monte Carlo dropout in the PINN framework, enabling estimation of prediction variability across temporal states and providing a potential foundation for future confidence-weighted temporal regularisation.

Third, although the proposed surrogates substantially reduce computation, careful attention is still required for boundary-condition specification, fibre architecture assumptions, and uncertainty propagation from image processing to biomechanics outputs, especially for clinical translation.

7.3 Future directions

Future work can build on this thesis in four directions. (i) Extending 3D+t tracking toward fully pre-trained deployment on larger multi-centre datasets to further reduce runtime and improve robustness under domain shift. (ii) Incorporating stronger physiological constraints on V-t (e.g. Windkessel-informed priors, explicit isovolumetric modelling, or uncertainty-aware temporal constraints) to improve haemodynamic fidelity when image-derived volumes are over-smoothed. (iii) Expanding surrogate outputs beyond displacement and pressure to richer biomechanical quantities (e.g. stress/strain distributions and clinically interpretable indices) with calibrated uncertainty estimates and improved identifiability under limited supervision. (iv) Strengthening clinical applicability through systematic validation against broader cohorts, alternative imaging protocols, and clinically measured haemodynamic endpoints, enabling more reliable deployment as a rapid, patient-adaptive modelling tool.

7.4 Closing remarks

Overall, this thesis contributes a coherent set of motion-consistent registration and multi-stage surrogate modelling methods that collectively move image-driven LV biomechanics toward fast, scalable, and physiologically grounded full-cycle simulation, while also identifying the remaining bottlenecks, particularly data limitations, input-signal fidelity, and clinical validation, that must be addressed for routine clinical translation.

Contents

8.1 Helix-Angle Definition for IMC-PINN-FE II (Biomechanical Parameter Estimation)	129
8.2 Default FEA Parameters in <i>IMC-PINN-FE</i>	130
8.3 Weights of Loss Components in <i>IMC-PINN-FE</i> III	130
8.4 Evaluated factors in <i>IMC-PINN-FE</i>	130
8.5 Sensitivity and Error Analysis of <i>IMC-PINN-FE</i>	132
8.5.1 Number of Temporal Frames	132
8.5.2 Robustness of the 1/3 Diastolic Unloading Reference	133
8.5.3 Cavity Volume Perturbation	135
8.5.4 Sensitivity of Imaging Noise	136
8.5.5 Loss Component Comparison	138
8.5.6 Reintroducing Isovolumetric Phases via a Windkessel-Constrained Volume–Time Curve	138
8.6 Data Used for Pre-trained Surrogate Comparison in <i>IMC-PINN-FE</i>	140

8.1 Helix-Angle Definition for IMC-PINN-FE II (Biomechanical Parameter Estimation)

Table 8.1 summarises the prescribed endocardial and epicardial helix angle settings used in this study for three model categories. The adult configuration was adopted from the Shavik setting [70], the fetal configuration was derived from Laura’s setting [128], and the pig configuration was based on Zheng’s setting [126]. In all configurations, the helix angles were assumed to vary linearly across the myocardial wall.

Table 8.1: Experimental endocardial and epicardial prescribed helix angle settings in adult, fetal, and pig models.

Case	Endo	Epi
Adult1	60	-60
Adult2	60	-60
Adult3	60	-60
Fetal1	71	-52
Fetal2	71	-52
Fetal3	71	-52
Pig-healthy1	12.5	-52.5
Pig-healthy2	15	-55
Pig-LVH1	0	-50
Pig-LVH2	-5	-55

8.2 Default FEA Parameters in *IMC-PINN-FE*

Table 8.2 provided FEA parameters used in our study for three different models. The adult configuration followed the Shavik setting [70], the fetal configuration was derived from Laura’s setting [128], and the pig configuration was based on Zheng’s setting [126].

8

8.3 Weights of Loss Components in *IMC-PINN-FE* III

Table 8.3 summarized the loss component weights used for training the volume-constrained PINN-FEA solver, determined through empirical testing of various combinations.

8.4 Evaluated factors in *IMC-PINN-FE*

To quantify the agreement between the predicted and reference P-V loops, we used three metrics, including the coefficient of determination R^2 for the entire loop, and the relative errors of end diastolic pressure (EDP-RE) and peak systolic pressure (PSP-RE). Let p_i and $\hat{p}_i$ denote the reference and predicted pressures at the i -th sampled time point, and let $\bar{p}$ be the mean of the reference pressures. The overall loop-fitting quality R^2 was measured by:

$$R^2 = 1 - \frac{\sum_i (p_i - \hat{p}_i)^2}{\sum_i (p_i - \bar{p})^2}. \quad (8.1)$$

Table 8.2: Default FEA parameters for adult, fetal, and pig models.

Parameter	Description	Adult	Fetus	Pig-healthy	Pig-LVH
b_{ff}	Material parameter	29.9	29.9	29.9	29.9
b_{xx}	Material parameter	13.3	13.3	13.3	13.3
b_{fx}	Material parameter	26.6	26.6	26.6	26.6
Ca_0	Peak intracellular calcium concentration (μM)	4.35	4.35	4.35	4.35
$(Ca_0)_{\max}$	Maximum peak intracellular calcium concentration (μM)	4.35	4.35	4.35	4.35
B	Governs the shape of the peak isometric tension-sarcomere length relation (μm^{-1})	4.75	4.75	4.75	4.75
l_0	Sarcomere length at which no active tension develops (μm)	1.58	1.58	1.58	1.58
l_r	Relaxed sarcomere length (μm)	1.85	1.85	1.85	1.85
BCL	Cardiac cycle length (ms)	800	400	530	530
t_0	Time to peak tension (ms)	171	140.5	160	160
m	Gradient of linear relaxation duration with sarcomere length ($\text{ms}\cdot\mu\text{m}^{-1}$)	524	524	524	524
b	Time intercept of linear relaxation duration with sarcomere length (ms)	-800	-835	-835	-835

Table 8.3: Weights of components for Volume-constrained PINN-FEA solver

w_1	w_2	w_3	w_4	w_5	w_6
$\text{loss}_{\text{energyT}}$ (μJ) ²	$\text{loss}_{\text{volume}}$ (mL) ²	loss_{ED} (mmHg) ²	loss_{PS} (mmHg) ²	$\text{loss}_{\text{dia trend}}$	$\text{loss}_{\text{sys trend}}$
1	1e3	1e2	2	1e6	1e4

The relative errors in end diastolic and peak systolic pressures were defined as:

$$\text{EDP-RE}(\%) = 100 \times \frac{|\hat{p}_{\text{EDP}} - p_{\text{EDP}}|}{p_{\text{EDP}}}, \quad (8.2)$$

$$\text{PSP-RE}(\%) = 100 \times \frac{|\hat{p}_{\text{PSP}} - p_{\text{PSP}}|}{p_{\text{PSP}}}, \quad (8.3)$$

where p_{EDP} and p_{PSP} are the reference end diastolic and peak systolic pressures, and $\hat{p}_{\text{EDP}}$, $\hat{p}_{\text{PSP}}$ are their predicted values. Together, R^2 , EDP-RE and PSP-RE captured both the global similarity of the P-V loop and the accuracy of the key physiological pressure levels.

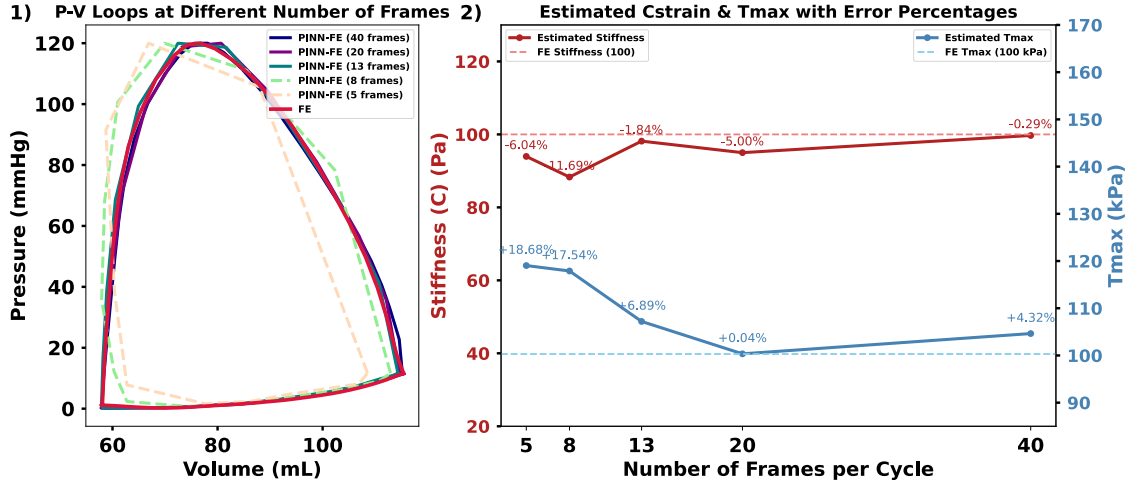


Figure 8.1: **Sensitivity Analysis of sampling intervals (Adult-3):** 1) P-V loops at different number of frames 2) Estimated C (red) and T_{max} (blue) with error percentages

8.5 Sensitivity and Error Analysis of *IMC-PINN-FE*

8.5.1 Number of Temporal Frames

To determine the optimal number of frames per cardiac cycle for image tracking (i.e., how temporal resolution affects the separability of different motion modes), we conducted a sensitivity analysis on the sampling interval, using the FEA validation case described in Chapter 4.3.1. Fig. 8.1 presents the results for the Adult-3 case, including (1) P-V loops under different numbers of frames (temporal resolutions) and (2) the estimated C (red) and T_{max} (blue) with corresponding error percentages. While *IMC-PINN-FE* maintained overall consistency across different frame numbers, a reduced number of frames increased the time interval per cycle, thereby lowering temporal resolution and causing inaccuracies in detecting key cardiac phases such as PS and ED. This resulted in deviations in the P-V relationship, particularly near PS.

Numerical analysis showed that when the number of frames falls below 8, the error in C surpassed 5%, and the deviation in T_{max} exceeded 10%, affecting estimator accuracy and PINN-FEA solver performance. Conversely, increasing the number of frames beyond 20 yielded minimal accuracy improvements while increasing computational cost. To balance accuracy and efficiency, we typically selected 20 frames for both adults (cycle duration: 800 ms) and fetuses (cycle duration: 400 ms), ensuring robust phase detection, reliable estimator outputs, and stable PINN-FEA solutions.

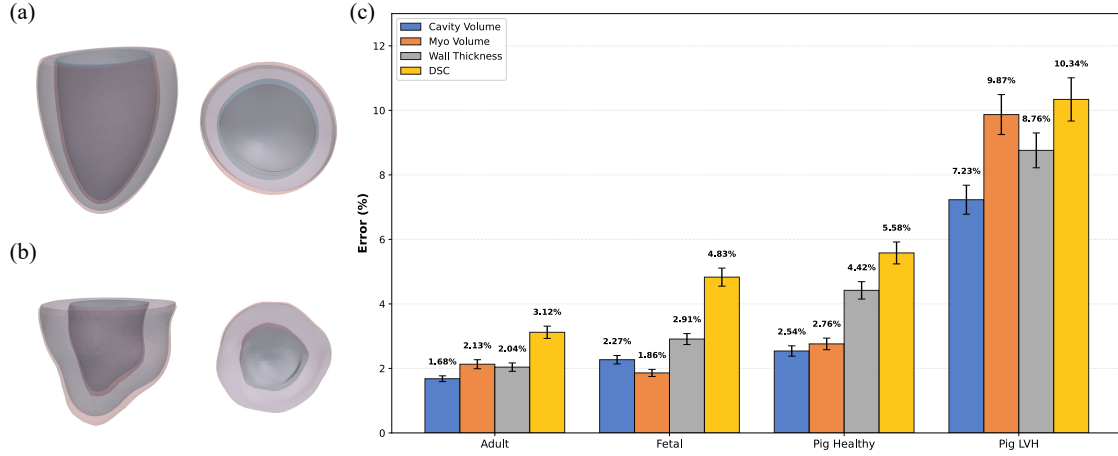


Figure 8.2: **Geometric comparison of the unloaded state between the one third diastolic assumption (blue surface) and that obtained from inverse FEA unloading (red surface)**, for (a) a representative healthy adult LV, and (b) a representative LVH pig case. A good agreement is found although a visible mismatch is at the apical region is found for (b). (c) Relative errors in cavity volume, myocardial volume, mean wall thickness and DSC between the unloaded LV from the two approaches, for 4 hearts (human adult, human fetal, healthy pig and LVH pig models). Small deviations were observed for most cases but there were moderate discrepancies in the LVH group.

Table 8.4: **Quantitative comparison of parameter back computation and P-V loop estimation when using the one third diastolic configuration as the unloaded state**, with the inverse FEA unloading used as reference. C-RE: relative error of myocardial stiffness, T_{max} -RE: maximal active tension, P-V R^2 : coefficient of determination for the P-V relation, EDP-RE: relative errors in end diastolic pressure, PSP-RE: peak systolic pressure.

Case	C-RE ↓ (%)	T_{max} -RE ↓ (%)	P-V R^2 ↑	EDP-RE ↓ (%)	PSP-RE ↓ (%)
Adult	1.3 ± 0.1	1.1 ± 0.2	0.99 ± 0.01	0.5 ± 0.1	0.4 ± 0.1
Fetus	2.2 ± 0.3	2.1 ± 0.4	0.99 ± 0.01	3.1 ± 0.4	2.5 ± 0.3
Pig-healthy	1.7 ± 0.4	1.9 ± 0.2	0.98 ± 0.01	7.9 ± 0.3	3.1 ± 0.3
Pig-LVH	4.5 ± 0.2	2.1 ± 0.6	0.98 ± 0.01	10.9 ± 1.1	4.0 ± 0.4

8.5.2 Robustness of the 1/3 Diastolic Unloading Reference

In this section, to evaluate the error introduced by taking the one-third diastolic configuration ($\Omega_0^{1/3}$) as the unloaded state, we compared it with the unloaded geometry obtained from inverse FEA unloading (Ω_0^{invFE}) as a reference. Inverse FEA unloading is a classical approach for computing a zero-pressure reference configuration [160], and in our implementation, the unloaded geometry was computed iteratively such that forward-loading FEA modelling could reverse the unloading and recover the target end-diastolic cavity volume and pressure.

Fig. 8.2 shows the comparison results. For the adult and fetal cases, the differences in cavity volume and mean wall thickness were mostly within 1%–3%, with DSC errors around 3%–4%.

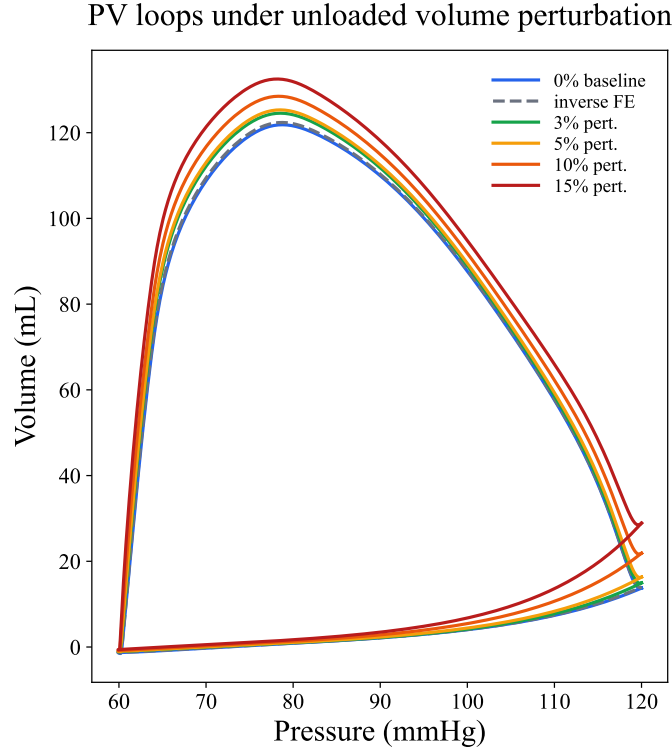


Figure 8.3: **P-V loops under unloaded volume perturbations (healthy adult case):** The blue solid line shows the 0% baseline, and the grey dashed line shows the inverse FE reference. The green, yellow, orange and dark red curves correspond to 3%, 5%, 10% and 15% increases in unloaded cavity volume, respectively, illustrating that small perturbations (3%–5%) have little effect on the loop shape, whereas larger perturbations (10%–15%) mainly shift the loop near end diastole and alter end diastolic pressure.

These discrepancies were slightly larger in the healthy pig models, and increased further in the LVH pig models, where the DSC difference reached approximately 10%. Consistent with these statistics, Fig. 3(a) shows a very close match between the human adult $\Omega_0^{1/3}$ and Ω_0^{invFE} LV geometries, while the LVH pig case exhibits a visible mismatch near the apical region, where the LV wall is slightly thicker for the inverse FEA approach.

On this basis, we then used both $\Omega_0^{1/3}$ and Ω_0^{invFE} within the full estimation pipeline, including parameter back computation and P-V loop estimation, to assess the impact of the unloaded-geometry assumption. Table 8.4 reports the relative errors of myocardial stiffness C and maximal active tension $T_{\max}$ for the $\Omega_0^{1/3}$ cases with respect to the Ω_0^{invFE} reference, together with the evaluation metrics R^2 , EDP-RE, and PSP-RE for the P-V relation. For the adult, fetal, and healthy pig models, the relative errors in C and $T_{\max}$ were approximately 1%–2%, whereas the LVH pig model showed a larger error of about 4.5% in C . The reconstructed P-V loops remained highly consistent, with R^2 values between 0.98 and 0.99, and the key pressure indices (EDP and PSP) were generally within 10% of the inverse FEA reference. The largest percentage deviation

occurred in EDP for the LVH pig, mainly due to its low absolute pressure level, and corresponded to an absolute difference of only a few millimetres of mercury.

Overall, these results indicate that using the one-third diastolic configuration as the unloaded state has a limited impact on parameter back computation and P-V loop estimation in most cases, although the discrepancy can be larger for certain LV geometries (e.g., LVH pigs).

8.5.3 Cavity Volume Perturbation

In this section, we further assessed the sensitivity of *IMC-PINN-FE* to errors in the choice of the unloaded cavity volume. The question was how much the back-computed stiffness C , maximal active tension $T_{\max}$, and the estimated P-V loop would change if the unloaded reference volume was slightly overestimated or underestimated along the diastolic filling trajectory.

We took the one-third diastolic configuration as the 0% baseline unloaded geometry, with LV cavity volume V_0 . Along the diastolic V-t curve of the same cardiac cycle, we then used interpolation to select time points at which the cavity volume was approximately $1.03V_0$, $1.05V_0$, $1.10V_0$, and $1.15V_0$, corresponding to 3%, 5%, 10%, and 15% volume perturbations. The geometries at these time points were treated as perturbed unloaded reference states. For each perturbation level, we used the corresponding geometry as the unloaded reference and repeated the full pipeline of parameter back computation and P-V loop estimation. All relative errors were computed with respect to the 0% unloaded baseline configuration.

Fig. 8.3 shows the P-V loops obtained under different unloaded volumes, and Table 5 summarises the corresponding quantitative metrics. For 3% and 5% volume perturbations, the relative errors in C and $T_{\max}$ remained below 3%, the coefficient of determination R^2 for the P-V relation stayed around 0.98, and the errors in EDP and PSP were both within 10%. When the volume perturbation increased to 10% and 15%, the errors in C , $T_{\max}$ and especially EDP started to grow markedly, with EDP reaching relative errors of up to about 81% (in the representative case shown in Fig. 8.3). In this case, the estimated EDP rose to roughly 30 mmHg, which is clearly outside the physiological range, whereas the PSP error remained relatively modest and the overall loop shape was still relatively stable.

Overall, these results indicated that shifting the unloaded reference volume by about 3–5% along the same diastolic trajectory mainly introduced a mild rescaling of the volume axis, without altering the dominant contraction and relaxation behaviours. Within this range, the *IMC-PINN-*

Table 8.5: **Sensitivity of parameter back-computation and P-V estimation to errors in the assumed unloaded-state cavity volume**, compared to the 0% unloaded-volume baseline. C-RE: relative error of myocardial stiffness; $T_{\max}$ -RE: maximal active tension; P-V R^2 : coefficient of determination for the P-V relation; EDP-RE: relative error of end-diastolic pressure; PSP-RE: relative error of peak systolic pressure.

Volume perturbation (%)	C-RE ↓ (%)	$T_{\max}$ -RE ↓ (%)	P-V R^2 ↑	EDP-RE ↓ (%)	PSP-RE ↓ (%)
3%	0.8 ± 0.3	1.2 ± 0.5	0.98 ± 0.01	5.3 ± 0.2	2.5 ± 0.5
5%	1.9 ± 0.5	2.5 ± 0.9	0.98 ± 0.01	8.1 ± 0.3	2.7 ± 0.3
10%	4.9 ± 1.4	6.8 ± 2.0	0.98 ± 0.01	33.2 ± 1.9	5.8 ± 1.5
15%	10.1 ± 3.3	9.2 ± 3.1	0.97 ± 0.01	81.1 ± 2.5	8.1 ± 1.1

FE model remained robust, with small errors in the back-computed stiffness and maximal tension and P-V loops that were consistent with the baseline and inverse FEA references. Only when the unloaded volume deviated by more than about 10% did the estimates of end-diastolic pressure become highly sensitive, whereas PSP and the global P-V pattern were still reasonably preserved.

8.5.4 Sensitivity of Imaging Noise

In practical applications, the LV geometry and motion were derived from medical images that were inevitably affected by imaging noise and tracking errors. To assess how such imperfections propagated through the POD representation and the *IMC-PINN-FE* inference, we performed a synthetic imaging-noise experiment. Starting from the original medical image sequences, we first normalised the image intensities to the range $[0, 1]$, and then added zero-mean Gaussian noise directly to the images before contour extraction. The noise amplitude was controlled by the standard deviation σ of the Gaussian perturbation and was expressed as a percentage of the normalised intensity. We explored several noise levels up to 50%. The noisy images were then processed by the same 3D tracking pipeline as in the main experiments, followed by POD decomposition and the full *IMC-PINN-FE* parameter back computation and P-V loop estimation.

We first examined the effect of image noise on the POD motion modes, quantifying the DSC between POD modes from noisy images and those from noise-free images (Fig. 8.4). The leading POD modes (1–4) remained highly stable up to the 20% noise level, with DSC close to 1. Noticeable degradation appeared mainly in higher-order modes (5–6), and only at higher noise levels above 30%. This indicated that image noise perturbed local high-frequency details but had a limited effect on global contraction and relaxation patterns.

We then quantified how these perturbations affected the *IMC-PINN-FE* parameter estimations and P-V predictions (Table 8.6). At the 5%, 12% and 20% noise levels, the relative errors in C

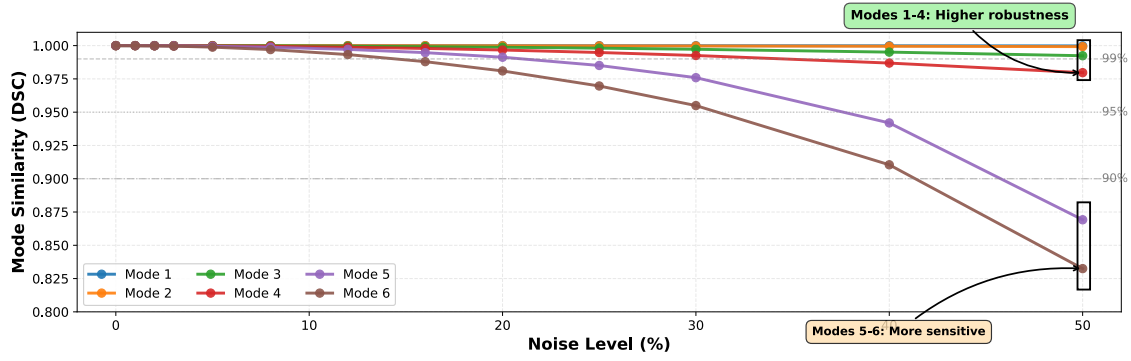


Figure 8.4: **Sensitivity of POD motion modes to imaging noise level.** Low-order modes (1–4) retain DSC values close to 1 and show high robustness even at moderate noise levels, whereas higher-order modes (5–6) are more sensitive and exhibit noticeable degradation only at the extreme 50% noise level.

Table 8.6: **Sensitivity of *IMC-PINN-FE* parameter estimation and P-V estimation to imaging noise.** Relative errors in myocardial stiffness (C-RE) and maximal active tension ($T_{\max}$ -RE), the coefficient of determination for the P-V relation (P-V R^2), and relative errors in end diastolic pressure (EDP-RE) and peak systolic pressure (PSP-RE) are reported for different noise levels, computed with respect to the noise-free reconstruction.

Noise level (%)	C-RE ↓ (%)	$T_{\max}$ -RE ↓ (%)	P-V R^2 ↑	EDP-RE ↓ (%)	PSP-RE ↓ (%)
5%	1.8 ± 0.7	0.8 ± 0.1	0.99 ± 0.01	2.3 ± 0.1	3.5 ± 0.2
12%	1.4 ± 0.3	0.5 ± 0.2	0.99 ± 0.01	3.1 ± 0.3	2.7 ± 0.4
20%	2.9 ± 0.7	1.4 ± 0.2	0.98 ± 0.01	8.2 ± 1.1	5.9 ± 0.5
50%	7.1 ± 1.3	8.4 ± 1.1	0.98 ± 0.01	31.1 ± 5.5	4.2 ± 0.8

and $T_{\max}$ remained below about 3%, R^2 for the P-V relation stayed around 0.99, and the errors in EDP and PSP were both within 10%. Only at the extreme 50% noise level did the relative errors in C and $T_{\max}$ rise to around 7–8%, and the EDP error increase to roughly 30%, while the PSP error remained below about 10% and the overall P-V loop shape was still relatively stable.

Overall, these results indicated that *IMC-PINN-FE* was robust to moderate levels of image noise, expected to be in the range of 5–20% noise. With unlikely severe noise, errors in EDP became large, whereas PSP and the global P-V pattern were still reasonably preserved. This robustness was consistent with the structure of the image-processing pipeline: the Fourier-based temporal regularisation and the truncation of POD modes acted as intrinsic low-pass filtering and smoothing steps, and the low-order POD modes were primarily governed by the global ventricular shape and contraction–relaxation pattern rather than by high-frequency local fluctuations.

8.5.5 Loss Component Comparison

Fig. 8.5 presents a comparative analysis of key losses across adult and fetal cases. Specifically, subplots 1 and 4 show the total energy loss ($\text{Loss}_{\text{energyT}}$), subplots 2 and 5 report the relative pressure loss at peak systole ($\sqrt{\text{Loss}_{\text{PS}}}/P_{\text{GT}}$), and subplots 3 and 6 illustrate the average relative volume loss ($\sqrt{\text{Loss}_{\text{volume}}}/V_{\text{GT}}$). Three methods are compared: *IMC-PINN-FE*, traditional volume-constrained FEA modelling, and Fixed Tracking. The metrics were computed using the same FEA validation cases used throughout this study.

The results clearly demonstrate that Fixed Tracking failed to balance pressure and energy constraints, leading to impractically large peak systolic pressure losses, often exceeding 60%, which undermined its physical reliability. This inconsistency was also reflected in the P-V loops (see Fig. 8.5 in the main text), where Fixed Tracking deviated substantially near peak systole. In contrast, *IMC-PINN-FE* achieved energy and volume losses comparable to the FEA baseline, indicating accurate enforcement of physiological constraints. Notably, *IMC-PINN-FE* not only matched FEA-level loss performance but also generated deformation patterns that were more consistent with the biological image sequences. This highlights the superiority of our method in combining physical consistency with anatomically realistic motion tracking, particularly under data-constrained or complex anatomical conditions.

8.5.6 Reintroducing Isovolumetric Phases via a Windkessel-Constrained Volume–Time Curve

To further examine whether isovolumetric phases can be recovered within our framework when the driving V-t curve itself already contains near-constant-volume segments, we performed an additional FEA experiment on a representative healthy adult LV case. First, we coupled the LV FEA model to a simplified three-element Windkessel afterload and ran a closed-loop simulation under physiologically reasonable loading conditions. From this coupled simulation, we extracted the chamber V-t trajectory, which showed a smooth profile with short near-flat intervals of almost constant volume around end-systole and end-diastole, corresponding to isovolumetric contraction and relaxation (Fig. 8.6, left).

We then used this Windkessel-derived V-t as the prescribed boundary condition to drive *IMC-PINN-FE* on the same LV geometry and with the same passive stiffness and active-tension time course. The resulting P-V loop predicted by *IMC-PINN-FE* was compared against the reference

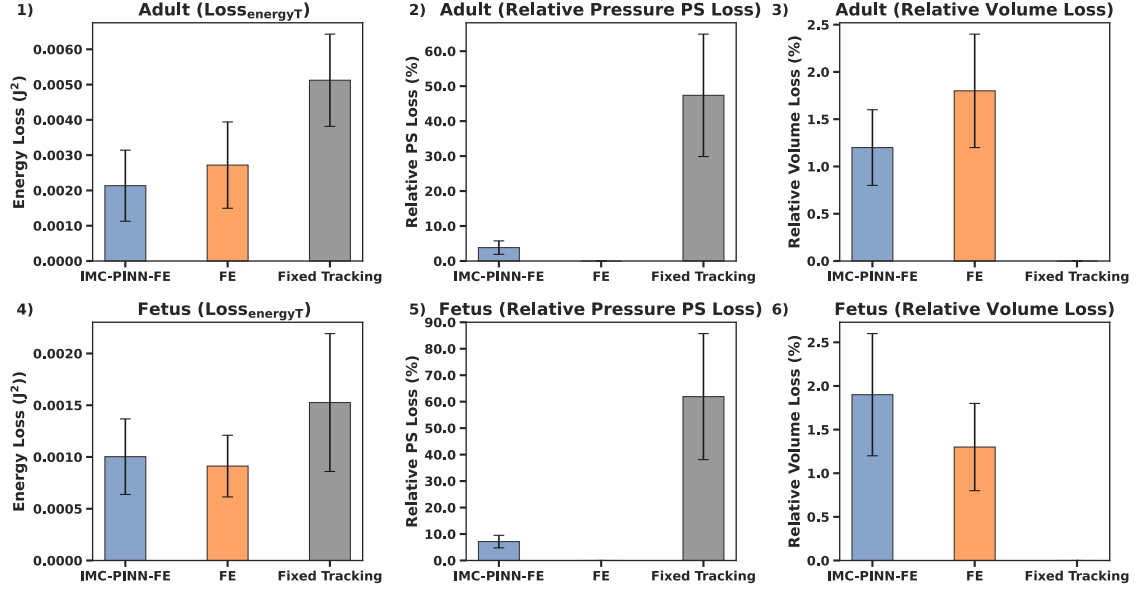


Figure 8.5: **Loss comparison across adult (1–3) and fetal (4–6) cases.** Subplots 1&4 show the total energy loss ($\text{Loss}_{\text{energyT}}$), subplots 2&5 show the relative pressure loss at peak systole ($\sqrt{\text{Loss}_{\text{PS}}}/P_{\text{GT}}$), and subplots 3&6 show the average relative volume loss ($\sqrt{\text{Loss}_{\text{volume}}}/V_{\text{GT}}$). *IMC-PINN-FE* consistently achieves low physics losses ($\text{Loss}_{\text{energyT}}$ and Loss_{PS}), comparable to the FEA baseline and substantially lower than Fixed Tracking.

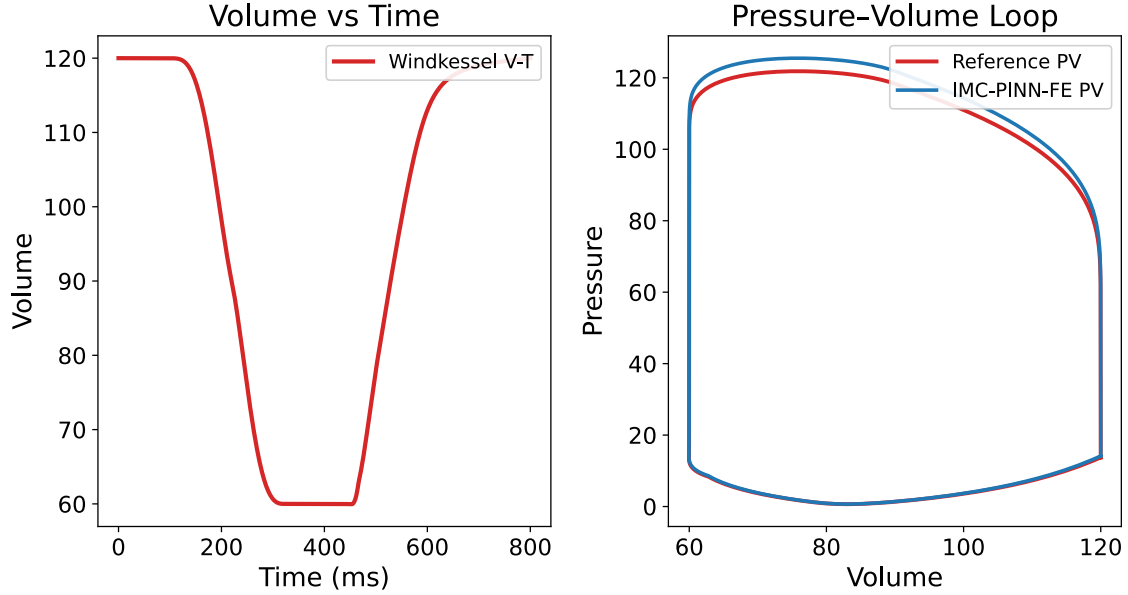


Figure 8.6: **Recovering isovolumetric phases via a Windkessel-constrained volume–time curve.** Left: chamber volume–time trajectory for a representative healthy adult LV obtained from an FEA–Windkessel closed-loop simulation, showing a smooth profile with short near-constant-volume segments around end-systole and end-diastole that correspond to isovolumetric contraction and relaxation. Right: comparison between the reference FEA–Windkessel P-V loop and the loop reconstructed by *IMC-PINN-FE* when driven by the same Windkessel-derived V-t curve, illustrating close agreement in overall loop shape, EDP and PSP, while clearly preserving the isovolumetric segments.

P-V loop directly obtained from the FEA–Windkessel simulation (Fig. 8.6, right). The two loops exhibit very close agreement in overall shape and key haemodynamic quantities, including EDP, PSP (slightly higher in the *IMC-PINN-FE* result) and stroke volume, while both clearly display near-vertical segments corresponding to isovolumetric contraction and relaxation.

This experiment indicated that, when provided with a physiologically plausible V-t curve that already contains well-defined isovolumetric stages, *IMC-PINN-FE* is able to preserve these features and reproduce the associated P-V loop with high fidelity. In other words, the current limitation regarding isovolumetric phases does not stem from the PINN–FEA coupling itself, but primarily from the smoothing inherent in image-derived volume–time curves. This suggests that, if the V-t input is appropriately constrained or post-processed (e.g. via Windkessel coupling or additional temporal constraints), the framework has the potential to recover more physiologically realistic isovolumetric behaviour without modifying the underlying myocardial mechanics.

8.6 Data Used for Pre-trained Surrogate Comparison in *IMC-PINN-FE*

In the pre-trained comparison experiments, all surrogate models were trained and evaluated on the same synthetic LV dataset. This dataset was generated by sampling from the publicly available PCA statistical shape model of the Cardiac Atlas Project [154], truncating each PCA coefficient to the range $[-2, 2]$ (covering approximately 95% of population variability), and meshing each sampled ED shape into a tetrahedral FEA model using Gmsh (typically 400–1000 nodes and 1000–2000 elements), with case-by-case anatomical quality control. A total of 67 LV geometries were used, providing more than one million pressure–volume–time (P-V- t) mapping states. To ensure a fair comparison, the testing set was kept identical to that used for the *IMC-PINN-FE* experiments, so that the pre-trained surrogates and the PINN-based method were evaluated on the same held-out cases.

Bibliography

- [1] T. D. Nguyen, O. E. Kadri, and R. S. Voronov, “An introductory overview of image-based computational modeling in personalized cardiovascular medicine,” *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 529 365, 2020.
- [2] V. Wang, P. Nielsen, and M. Nash, “Image-based predictive modeling of heart mechanics,” *Annual review of biomedical engineering*, vol. 17, no. 1, pp. 351–383, 2015.
- [3] L. Green, W. X. Chan, A. Tulzer, G. Tulzer, and C. H. Yap, “Myocardial biomechanical effects of fetal aortic valvuloplasty,” *Biomechanics and Modeling in Mechanobiology*, vol. 23, no. 5, pp. 1433–1448, 2024.
- [4] P. Hoang and J. Guccione, “Finite element modeling in left ventricular cardiac biomechanics: From computational tool to clinical practice,” *Bioengineering*, vol. 12, no. 9, p. 913, 2025.
- [5] H. Gao, H. Wang, C. Berry, X. Luo, and B. E. Griffith, “Quasi-static image-based immersed boundary-finite element model of left ventricle under diastolic loading,” *International journal for numerical methods in biomedical engineering*, vol. 30, no. 11, pp. 1199–1222, 2014.
- [6] D. E. Hurtado, N. Villarroel, J. Retamal, G. Buggedo, and A. Bruhn, “Improving the accuracy of registration-based biomechanical analysis: A finite element approach to lung regional strain quantification,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 2, pp. 580–588, 2015.
- [7] W. Li, “Biomechanics of infarcted left ventricle: A review of modelling,” *Biomedical Engineering Letters*, vol. 10, no. 3, pp. 387–417, 2020.
- [8] A. Dokuchaev et al., “Combination of personalized computational modeling and machine learning for optimization of left ventricular pacing site in cardiac resynchronization therapy,” *Frontiers in Physiology*, vol. 14, p. 1 162 520, 2023.
- [9] C. Rodero, T. M. Baptiste, R. K. Barrows, A. Lewalle, S. A. Niederer, and M. Strocchi, “Advancing clinical translation of cardiac biomechanics models: A comprehensive review, applications and future pathways,” *Frontiers in physics*, vol. 11, p. 1 306 210, 2023.

- [10] L. Cicci, S. Fresca, A. Manzoni, and A. Quarteroni, “Efficient approximation of cardiac mechanics through reduced-order modeling with deep learning-based operator approximation,” *International Journal for Numerical Methods in Biomedical Engineering*, vol. 40, no. 1, e3783, 2024.
- [11] R. Avazmohammadi, J. S. Soares, D. S. Li, S. S. Raut, R. C. Gorman, and M. S. Sacks, “A contemporary look at biomechanical models of myocardium,” *Annual review of biomedical engineering*, vol. 21, no. 1, pp. 417–442, 2019.
- [12] L. Marx, J. A. Niestrawska, M. A. Gsell, F. Caforio, G. Plank, and C. M. Augustin, “Robust and efficient fixed-point algorithm for the inverse elastostatic problem to identify myocardial passive material parameters and the unloaded reference configuration,” *Journal of computational physics*, vol. 463, p. 111 266, 2022.
- [13] J. H. Bracamonte, S. K. Saunders, J. S. Wilson, U. T. Truong, and J. S. Soares, “Patient-specific inverse modeling of in vivo cardiovascular mechanics with medical image-derived kinematics as input data: Concepts, methods, and applications,” *Applied Sciences*, vol. 12, no. 8, p. 3954, 2022.
- [14] C. Capelli et al., “Patient-specific simulations for planning treatment in congenital heart disease,” *Interface Focus*, vol. 8, no. 1, p. 20 170 021, 2018.
- [15] S. A. Niederer et al., “Creation and application of virtual patient cohorts of heart models,” *Philosophical Transactions of the Royal Society A*, vol. 378, no. 2173, p. 20 190 558, 2020.
- [16] C. Chen et al., “Deep learning for cardiac image segmentation: A review,” *Frontiers in cardiovascular medicine*, vol. 7, p. 25, 2020.
- [17] A. Sotiras, C. Davatzikos, and N. Paragios, “Deformable medical image registration: A survey,” *IEEE transactions on medical imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.
- [18] D. Dalton, H. Gao, and D. Husmeier, “Emulation of cardiac mechanics using graph neural networks,” *Computer Methods in Applied Mechanics and Engineering*, vol. 401, p. 115 645, 2022.
- [19] D. Dalton, D. Husmeier, and H. Gao, “Physics-informed graph neural network emulation of soft-tissue mechanics,” *Computer Methods in Applied Mechanics and Engineering*, vol. 417, p. 116 351, 2023.
- [20] L. Shi, Y. Chen, and V. Vedula, “Heartsimsage: Attention-enhanced graph neural networks for accelerating cardiac mechanics modeling,” *arXiv preprint arXiv:2504.18968*, 2025.

- [21] S. Motiwale, W. Zhang, R. Feldmeier, and M. S. Sacks, “A neural network finite element approach for high speed cardiac mechanics simulations,” *Computer Methods in Applied Mechanics and Engineering*, vol. 427, p. 117 060, 2024.
- [22] S. Buoso, T. Joyce, and S. Kozerke, “Personalising left-ventricular biophysical models of the heart using parametric physics-informed neural networks,” *Medical Image Analysis*, vol. 71, p. 102 066, 2021.
- [23] S. Mu, W. X. Chan, and C. H. Yap, “Imc-pinn-fe: A physics-informed neural network for patient-specific left ventricular finite element modeling with image motion consistency and biomechanical parameter estimation,” *arXiv preprint arXiv:2506.20696*, 2025.
- [24] S. Mu, W. X. Chan, and C. H. Yap, “Heartunloadnet: A cycle-consistent graph network with reduced supervision for predicting unloaded cardiac geometry from diastolic states,” *Computer Methods and Programs in Biomedicine*, p. 109 243, 2026.
- [25] S. Mu, W. X. Chan, and C. H. Yap, “A cycle-consistent graph surrogate for full-cycle left ventricular myocardial biomechanics,” *arXiv preprint arXiv:2602.06884*, 2026.
- [26] S. Mu, W. Chan, L. Song, and C. H. Yap, “An unsupervised deep learning framework for full cycle cardiac motion tracking in cross-modality images,” in *29th Congress of the European Society of Biomechanics (ESBiomech 2024): Abstracts*, 2024.
- [27] S. Mu, W. X. Chan, and C. H. Yap, “A PINN network for patient-specific left ventricular finite element modeling with image motion-consistency,” in *30th Congress of the European Society of Biomechanics (ESBiomech 2025): Abstracts*, 2025.
- [28] S. Mu, W. X. Chan, and C. H. Yap, “A physics-informed neural network for patient-specific left ventricular finite element modeling with image motion-consistency and myocardial stiffness and active tension estimation,” in *ASME SB3C Summer Bioengineering Conference (SBC 2025) Proceedings Book: Abstracts*, 2025.
- [29] L. Li et al., “The detection of viable myocardium by low-dose dobutamine stress speckle tracking echocardiography in patients with old myocardial infarction,” Dec. 2021.
- [30] W. Yu, P. Yan, A. J. Sinusas, K. Thiele, and J. S. Duncan, “Towards pointwise motion tracking in echocardiographic image sequences—comparing the reliability of different features for speckle tracking,” *Medical image analysis*, vol. 10, no. 4, pp. 495–508, 2006.
- [31] B. K. Horn and B. G. Schunck, “Determining optical flow,” *Artificial intelligence*, vol. 17, no. 1-3, pp. 185–203, 1981.

- [32] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *IJCAI'81: 7th international joint conference on Artificial intelligence*, vol. 2, 1981, pp. 674–679.
- [33] E. D. Angelini and O. Gerard, "Review of myocardial motion estimation methods from optical flow tracking on ultrasound data," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2006, pp. 1537–1540.
- [34] P. Torkashvand, H. Behnam, and Z. A. Sani, "Modified optical flow technique for cardiac motions analysis in echocardiography images," *Journal of Medical Signals & Sensors*, vol. 2, no. 3, pp. 121–127, 2012.
- [35] J. C. McEachen and J. S. Duncan, "Shape-based tracking of left ventricular wall motion," *IEEE Transactions on Medical Imaging*, vol. 16, no. 3, pp. 270–283, 2002.
- [36] T. Makela et al., "A review of cardiac image registration methods," *IEEE Transactions on medical imaging*, vol. 21, no. 9, pp. 1011–1021, 2002.
- [37] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Non-rigid registration using free-form deformations: Application to breast mr images," *IEEE transactions on medical imaging*, vol. 18, no. 8, pp. 712–721, 2002.
- [38] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, S61–S72, 2009.
- [39] M. J. Ledesma-Carbayo et al., "Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation," *IEEE transactions on medical imaging*, vol. 24, no. 9, pp. 1113–1126, 2005.
- [40] M. De Craene et al., "Temporal diffeomorphic free-form deformation: Application to motion and strain estimation from 3d echocardiography," *Medical image analysis*, vol. 16, no. 2, pp. 427–450, 2012.
- [41] H. Wiputra, W. X. Chan, Y. Y. Foo, S. Ho, and C. H. Yap, "Cardiac motion estimation from medical images: A regularisation framework applied on pairwise image registration displacement fields," *Scientific reports*, vol. 10, no. 1, p. 18 510, 2020.
- [42] W. X. Chan, Y. Zheng, H. Wiputra, H. L. Leo, and C. H. Yap, "Full cardiac cycle asynchronous temporal compounding of 3d echocardiography images," *Medical Image Analysis*, vol. 74, p. 102 229, 2021.

- [43] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, “An unsupervised learning model for deformable medical image registration,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9252–9260.
- [44] M. Jaderberg, K. Simonyan, A. Zisserman, et al., “Spatial transformer networks,” *Advances in neural information processing systems*, vol. 28, 2015.
- [45] E. Evain et al., “Motion estimation by deep learning in 2d echocardiography: Synthetic dataset and validation,” *IEEE transactions on medical imaging*, vol. 41, no. 8, pp. 1911–1924, 2022.
- [46] K. Ta, S. S. Ahn, A. Lu, J. C. Stendahl, A. J. Sinusas, and J. S. Duncan, “A semi-supervised joint learning approach to left ventricular segmentation and motion tracking in echocardiography,” in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2020, pp. 1734–1737.
- [47] H. Yu et al., “Foal: Fast online adaptive learning for cardiac motion estimation,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 4313–4323.
- [48] M. Ye et al., “Deeptag: An unsupervised deep learning method for motion tracking on cardiac tagging magnetic resonance images,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 7261–7271.
- [49] X. Song et al., “Cross-modal attention for multi-modal image registration,” *Medical Image Analysis*, vol. 82, p. 102612, 2022.
- [50] B. Heyde, D. Barbosa, P. Claus, F. Maes, and J. D’hooge, “Three-dimensional cardiac motion estimation based on non-rigid image registration using a novel transformation model adapted to the heart,” in *International Workshop on Statistical Atlases and Computational Models of the Heart*, Springer, 2012, pp. 142–150.
- [51] B. Heyde, D. Barbosa, P. Claus, F. Maes, and J. D’hooge, “Influence of the grid topology of free-form deformation models on the performance of 3d strain estimation in echocardiography,” in *International Conference on Functional Imaging and Modeling of the Heart*, Springer, 2013, pp. 308–315.
- [52] G. Piella et al., “Multiview diffeomorphic registration: Application to motion and strain estimation from 3d echocardiography,” *Medical image analysis*, vol. 17, no. 3, pp. 348–364, 2013.

- [53] M. Sermesant, C. Forest, X. Pennec, H. Delingette, and N. Ayache, “Deformable biomechanical models: Application to 4d cardiac image analysis,” *Medical image analysis*, vol. 7, no. 4, pp. 475–488, 2003.
- [54] Q. Meng et al., “Mulvimotion: Shape-aware 3d myocardial motion tracking from multi-view cardiac mri,” *IEEE transactions on medical imaging*, vol. 41, no. 8, pp. 1961–1974, 2022.
- [55] L. Teng, W. Feng, M. Zhu, and X. Li, “End-to-end unsupervised 4d cardiac motion tracking with spatiotemporal optical flow networks,” *IEEE Access*, 2024.
- [56] S. S. Ahn et al., “Co-attention spatial transformer network for unsupervised motion tracking and cardiac strain analysis in 3D echocardiography,” *Medical Image Analysis*, vol. 84, p. 102711, Feb. 2023, ISSN: 13618415. DOI: 10.1016/j.media.2022.102711. Accessed: May 17, 2024.
- [57] L. Green, W. X. Chan, I. Prakash, A. Tulzer, G. Tulzer, and C. H. Yap, “Pre-intervention myocardial stress is a good predictor of aortic valvuloplasty outcome for fetal critical aortic stenosis and evolving hlhs,” *The Journal of Physiology*, vol. 602, no. 4, pp. 663–681, 2024.
- [58] M. Ren, W. X. Chan, L. Green, M. L. Buist, and C. H. Yap, “Biventricular finite element modeling of the fetal heart in health and during critical aortic stenosis,” *Biomechanics and Modeling in Mechanobiology*, vol. 23, no. 4, pp. 1331–1345, 2024.
- [59] S. M. Shavik et al., “Computational modeling studies of the roles of left ventricular geometry, afterload, and muscle contractility on myocardial strains in heart failure with preserved ejection fraction,” *Journal of cardiovascular translational research*, vol. 14, no. 6, pp. 1131–1145, 2021.
- [60] G. Balaban et al., “In vivo estimation of elastic heterogeneity in an infarcted human heart,” *Biomechanics and modeling in mechanobiology*, vol. 17, no. 5, pp. 1317–1329, 2018.
- [61] H. Finsberg et al., “Computational quantification of patient-specific changes in ventricular dynamics associated with pulmonary hypertension,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 317, no. 6, H1363–H1375, 2019.
- [62] G. A. Holzapfel and R. W. Ogden, “Constitutive modelling of passive myocardium: A structurally based framework for material characterization,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 367, no. 1902, pp. 3445–3475, 2009.

- [63] S. M. Shavik, S. T. Wall, J. Sundnes, D. Burkhoff, and L. C. Lee, “Organ-level validation of a cross-bridge cycling descriptor in a left ventricular finite element model: Effects of ventricular loading on myocardial strains,” *Physiological reports*, vol. 5, no. 21, e13392, 2017.
- [64] S. Niederer, P. Hunter, and N. Smith, “A quantitative analysis of cardiac myocyte relaxation: A simulation study,” *Biophysical journal*, vol. 90, no. 5, pp. 1697–1722, 2006.
- [65] M. Strocchi et al., “A publicly available virtual cohort of four-chamber heart meshes for cardiac electro-mechanics simulations,” *PloS one*, vol. 15, no. 6, e0235145, 2020.
- [66] B. Baillargeon et al., “Human cardiac function simulator for the optimal design of a novel annuloplasty ring with a sub-valvular element for correction of ischemic mitral regurgitation,” *Cardiovascular engineering and technology*, vol. 6, no. 2, pp. 105–116, 2015.
- [67] J. Mojumder et al., “Computational analysis of ventricular mechanics in hypertrophic cardiomyopathy patients,” *Scientific reports*, vol. 13, no. 1, p. 958, 2023.
- [68] F. Mohammadi, S. M. Shontz, and C. A. Linte, “High-order cardiomyopathy human heart model and mesh generation,” in *2021 Computing in Cardiology (CinC)*, IEEE, vol. 48, 2021, pp. 1–4.
- [69] J. M. Guccione, A. D. McCulloch, and L. Waldman, “Passive material properties of intact ventricular myocardium determined from a cylindrical model,” 1991.
- [70] S. M. Shavik, Z. Jiang, S. Baek, and L. C. Lee, “High spatial resolution multi-organ finite element modeling of ventricular-arterial coupling,” *Frontiers in physiology*, vol. 9, p. 119, 2018.
- [71] S. Avril, “Inverse problems in the characterization of soft connective tissue: Perspective for reproduction system,” in *Biomechanics of the Female Reproductive System: Breast and Pelvic Organs*, Elsevier, 2023, pp. 115–138.
- [72] L. Fan, J. S. Choy, F. Raissi, G. S. Kassab, and L. C. Lee, “Optimization of cardiac resynchronization therapy based on a cardiac electromechanics-perfusion computational model,” *Computers in biology and medicine*, vol. 141, p. 105 050, 2022.
- [73] B. Baillargeon, N. Rebelo, D. D. Fox, R. L. Taylor, and E. Kuhl, “The living heart project: A robust and integrative simulator for human heart function,” *European Journal of Mechanics-A/Solids*, vol. 48, pp. 38–47, 2014.
- [74] J. Corral-Acero et al., “The ‘digital twin’ to enable the vision of precision cardiology,” *European heart journal*, vol. 41, no. 48, pp. 4556–4564, 2020.

- [75] L. Liang, M. Liu, C. Martin, and W. Sun, “A deep learning approach to estimate stress distribution: A fast and accurate surrogate of finite-element analysis,” *Journal of The Royal Society Interface*, vol. 15, no. 138, p. 20170844, 2018.
- [76] Y. Dabiri, A. Van der Velden, K. L. Sack, J. S. Choy, G. S. Kassab, and J. M. Guccione, “Prediction of left ventricular mechanics using machine learning,” *Frontiers in physics*, vol. 7, p. 117, 2019.
- [77] G. D. Maso Talou, T. P. Babarenda Gamage, M. Sagar, and M. P. Nash, “Deep learning over reduced intrinsic domains for efficient mechanics of the left ventricle,” *Frontiers in Physics*, vol. 8, p. 30, 2020.
- [78] Y. Dabiri, A. Van der Velden, K. L. Sack, J. S. Choy, J. M. Guccione, and G. S. Kassab, “Application of feed forward and recurrent neural networks in simulation of left ventricular mechanics,” *Scientific reports*, vol. 10, no. 1, p. 22298, 2020.
- [79] P. Di Achille, A. Harouni, S. Khamzin, O. Solovyova, J. J. Rice, and V. Gurev, “Gaussian process regressions for inverse problems and parameter searches in models of ventricular mechanics,” *Frontiers in physiology*, vol. 9, p. 1002, 2018.
- [80] U. Noè et al., “Gaussian process emulation to accelerate parameter estimation in a mechanical model of the left ventricle: A critical step towards clinical end-user relevance,” *Journal of the Royal Society Interface*, vol. 16, no. 156, p. 20190114, 2019.
- [81] V. Davies, U. Noè, A. Lazarus, H. Gao, et al., “Fast parameter inference in a biomechanical model of the left ventricle using statistical emulation,” *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 68, no. 5, pp. 1555–1576, 2019.
- [82] L. Cai, L. Ren, Y. Wang, W. Xie, G. Zhu, and H. Gao, “Surrogate models based on machine learning methods for parameter estimation of left ventricular myocardium,” *Royal Society open science*, vol. 8, no. 1, p. 201121, 2021.
- [83] H. Babaei et al., “A machine learning model to estimate myocardial stiffness from edpvr,” *Scientific Reports*, vol. 12, no. 1, p. 5433, 2022.
- [84] J. Chen et al., “Feaorta: A fully automated framework for finite element analysis of the aorta from 3d ct images,” *arXiv preprint arXiv:2510.06621*, 2025.
- [85] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, “Physics-informed machine learning,” *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, 2021.

- [86] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, “Scientific machine learning through physics-informed neural networks: Where we are and what’s next,” *Journal of Scientific Computing*, vol. 92, no. 3, p. 88, 2022.
- [87] S. Cai, Z. Mao, Z. Wang, M. Yin, and G. E. Karniadakis, “Physics-informed neural networks (pinns) for fluid mechanics: A review,” *Acta Mechanica Sinica*, vol. 37, no. 12, pp. 1727–1738, 2021.
- [88] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational physics*, vol. 378, pp. 686–707, 2019.
- [89] J. Garay, J. Dunstan, S. Uribe, and F. S. Costabal, “Physics-informed neural networks for parameter estimation in blood flow models,” *Computers in Biology and Medicine*, vol. 178, p. 108706, 2024.
- [90] P. A. López, H. Mella, S. Uribe, D. E. Hurtado, and F. S. Costabal, “WarpPINN: Cine-MR image registration with physics-informed neural networks,” *Medical Image Analysis*, vol. 89, p. 102925, Oct. 2023, ISSN: 13618415. DOI: 10.1016/j.media.2023.102925. Accessed: May 17, 2024.
- [91] X. Zhuang, “Challenges and methodologies of fully automatic whole heart segmentation: A review,” *Journal of healthcare engineering*, vol. 4, no. 3, pp. 371–407, 2013.
- [92] X. Zhuang, K. S. Rhode, R. S. Razavi, D. J. Hawkes, and S. Ourselin, “A registration-based propagation framework for automatic whole heart segmentation of cardiac mri,” *IEEE transactions on medical imaging*, vol. 29, no. 9, pp. 1612–1625, 2010.
- [93] X. Zhuang and J. Shen, “Multi-scale patch and multi-modality atlases for whole heart segmentation of mri,” *Medical image analysis*, vol. 31, pp. 77–87, 2016.
- [94] M. Ye, D. Yang, Q. Huang, M. Kanski, L. Axel, and D. N. Metaxas, “Sequencemorph: A unified unsupervised learning framework for motion tracking on cardiac image sequences,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.
- [95] Z. B. Popović et al., “Association between regional ventricular function and myocardial fibrosis in hypertrophic cardiomyopathy assessed by speckle tracking echocardiography and delayed hyperenhancement magnetic resonance imaging,” *Journal of the American Society of Echocardiography*, vol. 21, no. 12, pp. 1299–1305, 2008.

- [96] D. Seo, J. Ho, J. H. Traverse, J. Forder, and B. Vemuri, "Computing diffeomorphic paths with applications to cardiac motion analysis," in *4th MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, Citeseer, 2013, pp. 83–94.
- [97] C. W. Ong et al., "Biomechanics of human fetal hearts with critical aortic stenosis," *Annals of biomedical engineering*, vol. 49, pp. 1364–1379, 2021.
- [98] S. Mondillo et al., "Speckle-tracking echocardiography: A new technique for assessing myocardial function," *Journal of Ultrasound in Medicine*, vol. 30, no. 1, pp. 71–83, 2011.
- [99] N. Parajuli et al., "Flow network tracking for spatiotemporal and periodic point matching: Applied to cardiac motion analysis," *Medical image analysis*, vol. 55, pp. 116–135, 2019.
- [100] S. S. Ahn et al., "Co-attention spatial transformer network for unsupervised motion tracking and cardiac strain analysis in 3d echocardiography," *Medical image analysis*, vol. 84, p. 102711, 2023.
- [101] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "Voxelmorph: A learning framework for deformable medical image registration," *IEEE transactions on medical imaging*, vol. 38, no. 8, pp. 1788–1800, 2019.
- [102] K. Marstal, F. Berendsen, M. Staring, and S. Klein, "Simpleelastix: A user-friendly, multi-lingual library for medical image registration," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2016, pp. 134–142.
- [103] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. Pluim, "Elastix: A toolbox for intensity-based medical image registration," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 196–205, 2009.
- [104] S. Lee, G. Wolberg, and S. Shin, "Scattered data interpolation with multilevel b-splines," *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, no. 3, pp. 228–244, 1997. DOI: 10.1109/2945.620490.
- [105] O. Bernard et al., "Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is the problem solved?" *IEEE transactions on medical imaging*, vol. 37, no. 11, pp. 2514–2525, 2018.
- [106] S. Leclerc et al., "Deep learning for segmentation using an open large-scale dataset in 2d echocardiography," *IEEE transactions on medical imaging*, vol. 38, no. 9, pp. 2198–2210, 2019.

- [107] H. Wiputra et al., “Fluid mechanics of human fetal right ventricles from image-based computational fluid dynamics using 4d clinical ultrasound scans,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 311, no. 6, H1498–H1508, 2016.
- [108] O. U. Aydin et al., “On the usage of average hausdorff distance for segmentation performance assessment: Hidden error when used for ranking,” *European radiology experimental*, vol. 5, pp. 1–7, 2021.
- [109] M. K. Hasan, H. Zhu, G. Yang, and C. H. Yap, “A multi-scale learning of data-driven and anatomically constrained image registration for adult and fetal echo images,” *arXiv preprint arXiv:2309.00831*, 2023.
- [110] Y. Zhang et al., “A novel mri-based finite element modeling method for calculation of myocardial ischemia effect in patients with functional mitral regurgitation,” *Frontiers in physiology*, vol. 11, p. 158, 2020.
- [111] A. Salih, F. Hamandi, and T. Goswami, “Advancements in finite element modeling for cardiac device leads and 3d heart models,” *Bioengineering*, vol. 11, no. 6, p. 564, 2024.
- [112] M. Genet, C. T. Stoeck, C. von Deuster, L. C. Lee, and S. Kozerke, “Equilibrated warping: Finite element image registration with finite strain equilibrium gap regularization,” *Medical image analysis*, vol. 50, pp. 1–22, 2018.
- [113] X. Zhang et al., “Fast virtual stenting for thoracic endovascular aortic repair of aortic dissection using graph deep learning,” *IEEE Journal of Biomedical and Health Informatics*, 2025.
- [114] A. E. Ilesanmi, T. O. Ilesanmi, and B. O. Ajayi, “Reviewing 3d convolutional neural network approaches for medical image segmentation,” *Heliyon*, 2024.
- [115] I. Van Heerden, A. Fossey, and G. J. Booysen, “3d digital geometry designs for poland’s syndrome using magics and geomagic® freeform®,” *Rapid Prototyping Journal*, vol. 24, no. 1, pp. 229–236, 2018.
- [116] Y. Luo et al., *Explicit differentiable slicing and global deformation for cardiac mesh reconstruction*, 2024. arXiv: 2409.02070 [eess.IV]. [Online]. Available: <https://arxiv.org/abs/2409.02070>.
- [117] C. Geuzaine and J.-F. Remacle, “Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities,” *International journal for numerical methods in engineering*, vol. 79, no. 11, pp. 1309–1331, 2009.

- [118] J. F. Wenk et al., “First evidence of depressed contractility in the border zone of a human myocardial infarction,” *The Annals of thoracic surgery*, vol. 93, no. 4, pp. 1188–1193, 2012.
- [119] H. Gao, D. Carrick, C. Berry, B. E. Griffith, and X. Luo, “Dynamic finite-strain modelling of the human left ventricle in health and disease using an immersed boundary-finite element method,” *IMA journal of applied mathematics*, vol. 79, no. 5, pp. 978–1010, 2014.
- [120] C. Mancinelli, M. Livesu, and E. Puppo, “A comparison of methods for gradient field estimation on simplicial meshes,” *Computers & Graphics*, vol. 80, pp. 37–50, 2019.
- [121] S. Buoso, A. Manzoni, H. Alkadhi, A. Plass, A. Quarteroni, and V. Kurtcuoglu, “Reduced-order modeling of blood flow for noninvasive functional evaluation of coronary artery disease,” *Biomechanics and modeling in mechanobiology*, vol. 18, pp. 1867–1881, 2019.
- [122] A. Manzoni, A. Quarteroni, and G. Rozza, “Model reduction techniques for fast blood flow simulation in parametrized geometries,” *International journal for numerical methods in biomedical engineering*, vol. 28, no. 6-7, pp. 604–625, 2012.
- [123] A. Quarteroni and G. Rozza, “Numerical solution of parametrized navier–stokes equations by reduced basis methods,” *Numerical methods for partial differential equations: an international journal*, vol. 23, no. 4, pp. 923–948, 2007.
- [124] C. W. Rowley, “Model reduction for fluids, using balanced proper orthogonal decomposition,” *International Journal of Bifurcation and Chaos*, vol. 15, no. 03, pp. 997–1013, 2005.
- [125] D. Burkhoff, I. Mirsky, and H. Suga, “Assessment of systolic and diastolic ventricular properties via pressure-volume analysis: A guide for clinical, translational, and basic researchers,” *American Journal of Physiology-Heart and Circulatory Physiology*, 2005.
- [126] Y. Zheng et al., “Morphological, functional, and biomechanical progression of lv remodelling in a porcine model of hfpef,” *Journal of biomechanics*, vol. 144, p. 111 348, 2022.
- [127] T. Arts, R. S. Reneman, and P. C. Veenstra, “A model of the mechanics of the left ventricle,” *Annals of biomedical engineering*, vol. 7, pp. 299–318, 1979.
- [128] L. Green, W. X. Chan, M. Ren, C. N. Z. Mattar, L. C. Lee, and C. H. Yap, “The dependency of fetal left ventricular biomechanics function on myocardium helix angle configuration,” *Biomechanics and Modeling in Mechanobiology*, vol. 22, no. 2, pp. 629–643, 2023.
- [129] NVIDIA Modulus Team, *Nvidia modulus*, Computer Software, 2023. [Online]. Available: <https://github.com/NVIDIA/modulus/tree/main>.
- [130] N. O. Fowler, *Cardiac Diagnosis and Treatment*, 3rd ed. Philadelphia: JB Lippincott, 1980, p. 11.

- [131] Merck Manuals Professional Version. “Normal pressures in the heart and great vessels.” Accessed: 2025-05-11. [Online]. Available: <https://www.merckmanuals.com/professional/multimedia/table/normal-blood-pressure-in-the-heart-and-great-vessels>.
- [132] P. Johnson, D. Maxwell, M. Tynan, and L. Allan, “Intracardiac pressures in the human fetus,” *Heart*, vol. 84, no. 1, pp. 59–63, 2000.
- [133] V. Y. Wang, H. I. Lam, D. B. Ennis, B. R. Cowan, A. A. Young, and M. P. Nash, “Cardiac active contraction parameters estimated from magnetic resonance imaging,” in *Statistical Atlases and Computational Models of the Heart: First International Workshop, STACOM 2010, and Cardiac Electrophysiological Simulation Challenge, CESC 2010, Held in Conjunction with MICCAI 2010, Beijing, China, September 20, 2010. Proceedings 1*, Springer, 2010, pp. 194–203.
- [134] A. Müller, E. Kovacheva, S. Schuler, O. Dössel, and L. Baron, “Effects of local activation times on the tension development of human cardiomyocytes in a computational model,” *Current Directions in Biomedical Engineering*, vol. 4, no. 1, pp. 247–250, 2018.
- [135] H. Gao, B. Griffith, D. Carrick, C. McComb, C. Berry, and X. Luo, “Electro-mechanical modelling of in-vivo human left ventricle,” *Journal of Electrocardiology*, vol. 46, no. 4, e26, 2013.
- [136] A. W. Racca et al., “Contractile properties of developing human fetal cardiac muscle,” *The Journal of physiology*, vol. 594, no. 2, pp. 437–452, 2016.
- [137] T. Pfaff, M. Fortunato, A. Sanchez-Gonzalez, and P. Battaglia, “Learning mesh-based simulation with graph networks,” in *International conference on learning representations*, 2020.
- [138] W. Hamilton, Z. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” *Advances in neural information processing systems*, vol. 30, 2017.
- [139] H. Gao and S. Ji, “Graph u-nets,” in *international conference on machine learning*, PMLR, 2019, pp. 2083–2092.
- [140] A. Nikou et al., “Effects of using the unloaded configuration in predicting the in vivo diastolic properties of the heart,” *Computer methods in biomechanics and biomedical engineering*, vol. 19, no. 16, pp. 1714–1720, 2016.
- [141] M. Genet, L. C. Lee, B. Baillargeon, J. M. Guccione, and E. Kuhl, “Modeling pathologies of diastolic and systolic heart failure,” *Annals of biomedical engineering*, vol. 44, pp. 112–127, 2016.

- [142] Z. J. Wang et al., “Efficient estimation of load-free left ventricular geometry and passive myocardial properties using principal component analysis,” *International journal for numerical methods in biomedical engineering*, vol. 36, no. 3, e3313, 2020.
- [143] A. Krishnamurthy et al., “Patient-specific models of cardiac biomechanics,” *Journal of computational physics*, vol. 244, pp. 4–21, 2013.
- [144] M. J. Gonzales et al., “A three-dimensional finite element model of human atrial anatomy: New methods for cubic hermite meshes with extraordinary vertices,” *Medical image analysis*, vol. 17, no. 5, pp. 525–537, 2013.
- [145] L. Liang, M. Liu, C. Martin, and W. Sun, “A machine learning approach as a surrogate of finite element analysis-based inverse method to estimate the zero-pressure geometry of human thoracic aorta,” *International journal for numerical methods in biomedical engineering*, vol. 34, no. 8, e3103, 2018.
- [146] L. Liang, M. Liu, J. Elefteriades, and W. Sun, “Pytorch-fea: Autograd-enabled finite element analysis methods with applications for biomechanical analysis of human aorta,” *Computer methods and programs in biomedicine*, vol. 238, p. 107 616, 2023.
- [147] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” *arXiv preprint arXiv:1609.02907*, 2016.
- [148] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.
- [149] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, “Graph attention networks,” *arXiv preprint arXiv:1710.10903*, 2017.
- [150] Y. LeCun et al., “Backpropagation applied to handwritten zip code recognition,” *Neural computation*, vol. 1, no. 4, pp. 541–551, 1989.
- [151] A. Vaswani et al., “Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017.
- [152] J. L. Ba, J. R. Kiros, and G. E. Hinton, “Layer normalization,” *arXiv preprint arXiv:1607.06450*, 2016.
- [153] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

- [154] P. Medrano-Gracia et al., “Left ventricular shape variation in asymptomatic populations: The multi-ethnic study of atherosclerosis,” *Journal of Cardiovascular Magnetic Resonance*, vol. 16, no. 1, p. 56, 2014.
- [155] R. J. BOUCHARD, J. H. GAULT, and J. Ross Jr, “Evaluation of pulmonary arterial end-diastolic pressure as an estimate of left ventricular end-diastolic pressure in patients with normal and abnormal left ventricular performance,” *Circulation*, vol. 44, no. 6, pp. 1072–1079, 1971.
- [156] L. Geerts, P. Bovendeerd, K. Nicolay, and T. Arts, “Characterization of the normal cardiac myofiber field in goat measured with mr-diffusion tensor imaging,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 283, no. 1, H139–H145, 2002.
- [157] Y. Zheng et al., “Effects of myocardial sheetlet sliding on left ventricular function,” *Biomechanics and Modeling in Mechanobiology*, vol. 22, no. 4, pp. 1313–1332, 2023.
- [158] J. Whitaker et al., “Improved co-registration of ex-vivo and in-vivo cardiovascular magnetic resonance images using heart-specific flexible 3d printed acrylic scaffold combined with non-rigid registration,” *Journal of Cardiovascular Magnetic Resonance*, vol. 21, no. 1, p. 62, 2019.
- [159] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, “Pointnet++: Deep hierarchical feature learning on point sets in a metric space,” *arXiv preprint arXiv:1706.02413*, 2017.
- [160] H. Finsberg et al., “Efficient estimation of personalized biventricular mechanical function employing gradient-based optimization,” *International journal for numerical methods in biomedical engineering*, vol. 34, no. 7, e2982, 2018.
- [161] S. Brody, U. Alon, and E. Yahav, “How attentive are graph attention networks?” *arXiv preprint arXiv:2105.14491*, 2021.
- [162] S. Yun, M. Jeong, R. Kim, J. Kang, and H. J. Kim, “Graph transformer networks,” *Advances in neural information processing systems*, vol. 32, 2019.
- [163] M. Lin, Q. Chen, and S. Yan, “Network in network,” *arXiv preprint arXiv:1312.4400*, 2013.
- [164] L. E. Hudsmith, S. E. Petersen, J. M. Francis, M. D. Robson, and S. Neubauer, “Normal human left and right ventricular and left atrial dimensions using steady state free precession magnetic resonance imaging,” *Journal of cardiovascular magnetic resonance*, vol. 7, no. 5, pp. 775–782, 2005.
- [165] J. Walpot et al., “Left ventricular mid-diastolic wall thickness: Normal values for coronary ct angiography,” *Radiology: Cardiothoracic Imaging*, vol. 1, no. 5, e190034, 2019.