Supplementary Document for Volumetrically Consistent Implicit Atlas Learning via Neural Diffeomorphic Flow for Placenta MRI

A. Overview

This supplementary document provides additional technical details, experiments, and visualizations supporting our main paper. In Sec. B, we describe the MRI dataset and outline the preprocessing steps used to extract surfaces, generate signed-distance samples, and construct tetrahedral embeddings. We then provide implementation specifics, including network architectures, training procedures, and evaluation settings.

B. Experimental Details

B.1. Datasets and Preprocessing

Our primary dataset consists of 111 EPI BOLD MRI scans acquired from 78 pregnant subjects on a 3T Siemens Skyra scanner (single-shot GRE–EPI, 3 mm isotropic voxels, interleaved slice acquisition, TR = 5.8–8 s, TE = 32–36 ms, FA = 90°). The cohort includes 60 singleton and 18 monochorionic twin pregnancies. Among singleton cases, 33 subjects were scanned in two maternal positions (supine and left lateral), providing 66 distinct placenta shapes, while the remaining 27 subjects were scanned in a single position. Gestational ages range from 27 to 38 weeks, covering mid-to-late pregnancy (Fig. 1). We used 91 placentas for training and 21 placentas for testing, ensuring that no subjects overlap between sets when multiple scans were available for the same pregnancy. All placentas were manually segmented by an expert; small holes were filled, and only the largest connected component was retained before meshing. The resulting volumetric meshes contain 6708 ± 1813 tetrahedra and 2801 ± 731 surface triangles.

Preprocessing. For each segmentation volume, a surface mesh was first extracted using the Marching Cubes algorithm [3]. Meshes were normalized to the unit cube and reoriented such that the first principal axis aligned with the z-axis to ensure a consistent anatomical "upright" pose. Signed-distance samples were generated using depth-based nearest-surface queries. For each case, we sampled points near the surface, producing a mixture of positive (outside), negative (inside), and surface-normal samples (Fig. 2). In addition, a voxelized SDF grid at resolution was computed to provide spatially dense supervision. Tetrahedral embeddings were constructed from the cleaned volumetric meshes and used exclusively for volumetric regularization and geometric evaluation, not for supervision of reconstruction loss.

B.2. Architecture Details

Implicit Template Decoder. The implicit template $\mathcal{T}_\theta$ is parameterized as a MLP similar to DeepSDF/DIT [4, 8]. The network takes a 3D point coordinate as input and passes it through five 256-width layers with ReLU activations, weight normalization, and a dropout probability of 0.05. A skip connection is inserted at the 4th layer as in DeepSDF. The final layer uses a hyperbolic tangent activation to produce a bounded SDF value: $3 \to 256 \to 256 \to 256 \to 256 \to 256 \to 1$. This continuous implicit representation allows surfaces to be reconstructed at arbitrary resolution, independent of the resolution of the voxel grid segmentation mesh.

Deformation module. We adopt a Neural ODE–based warping module following Neural Mesh Flow [1, 7]. Given a point p_j^0 sampled from the instance's SDF, the deformation Φ_i is defined as a composition of four concatenated NODE blocks, each integrating a learned velocity field over $t \in [0, 1]$. Each velocity field is estimated by a fully-connected residual network with hidden width 512: $(3+256) \to 512 \to 512 \to 512 \to 3$, with ReLU activations in intermediate layers and tanh at the output to keep displace-

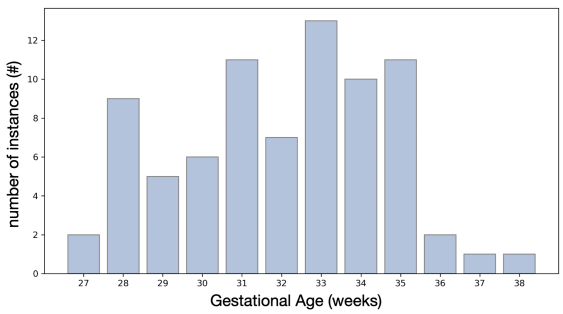

Figure 1. Distribution of gestational ages for our dataset of EPI BOLD placenta scans.

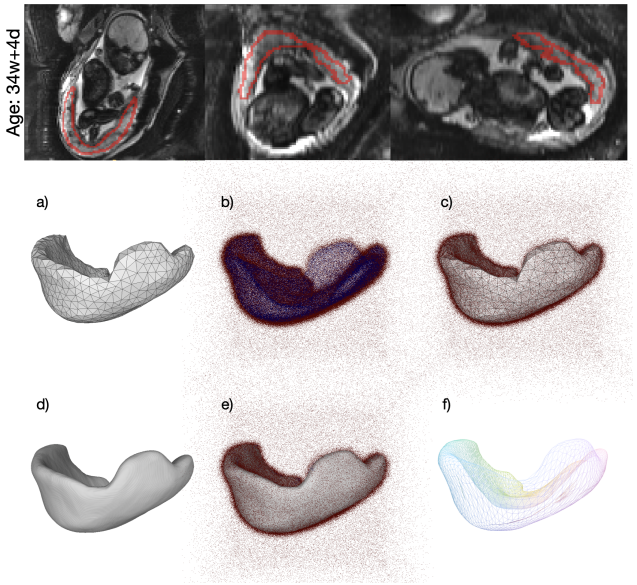

Figure 2. Preprocessing and model outputs for a representative instance with gestational age of 34w+4d (a) Initial surface extracted from the manual segmentation using Marching Cubes (b) Dense signed-distance samples generated using depth-based nearest-surface queries (red: outside; blue: inside)(c) Overlay of the SDF samples with the initial mesh; a random subset of 5k points is used for training (d) Reconstructed implicit surface produced by our template-conditioned SDF decoder; since the representation is continuous, the surface can be extracted at arbitrarily high resolution independently of the input mesh resolution (e) Overlay of the reconstructed surface with the input SDF samples, illustrating geometric consistency (f) Dense pointwise correspondences obtained from the learned diffeomorphic flow, mapping the subject into the template space.

ments within the normalized domain of $[-1, 1]$. The latent code c_i is concatenated to the input of every NODE block (Fig. 3).

Volumetric Embedding and Regularization. To enforce topological correctness and smooth volumetric deformation, we construct a tetrahedral embedding $\mathcal{M}_i = (V_i, K_i)$ for each instance from the cleaned segmentation mesh. During training, we sample 2,000 tetrahedra per instance and penalize negative Jacobian determinants of Φ_i using a hinge loss with $\epsilon_{\text{raito}} = 0.03$ and weight $\lambda_{\text{Jac}} = 0.5$. To promote smooth interior deformations, a biharmonic penalty is applied to 1,500 interior vertices using the symmetric normalization scheme and weight $\lambda_{\text{biH}} = 0.2$. These volumetric terms complement the SDF loss but do not supervise surface geometry.

B.3. Training and Inference Details

We jointly optimize network parameters and latent codes using Adam for 2000 epochs. The learning rate for the deformation and template networks is initialized at 5×10^{-4} and reduced by a factor of 0.5 every 500 epochs, while latent codes use a separate optimizer with learning rate 1×10^{-3}. Each batch contains four instances, and for every instance we randomly sample 5,000 SDF points. SDF values are clamped at a truncation band of 0.1. Latent codes are regularized by $\lambda \|\boldsymbol{c}_i\|_2^2$ with $\lambda = 10^{-4}$ and constrained to lie within a unit ℓ_2-ball. During inference, the latent codes are optimized for 2400 iterations using an Adam optimizer with a learning rate of 5×10^{-2}.

C. More Experiments

Diffeomorphism Analysis. We report the full distribution of diffeomorphic metrics including flip rate, absolute log-determinant, cycle-consistency error, and Laplacian energy sorted by increasing gestational age (GA). For each GA bin, values are averaged over all placentas within that interval, with shaded regions indicating variability across instances. Across all gestational ages, our method consistently achieves lower flip rates, smaller $|\text{LogDet}|$, reduced cycle error, and lower Laplacian energy compared to baseline methods, demonstrating more stable and smoother vol-

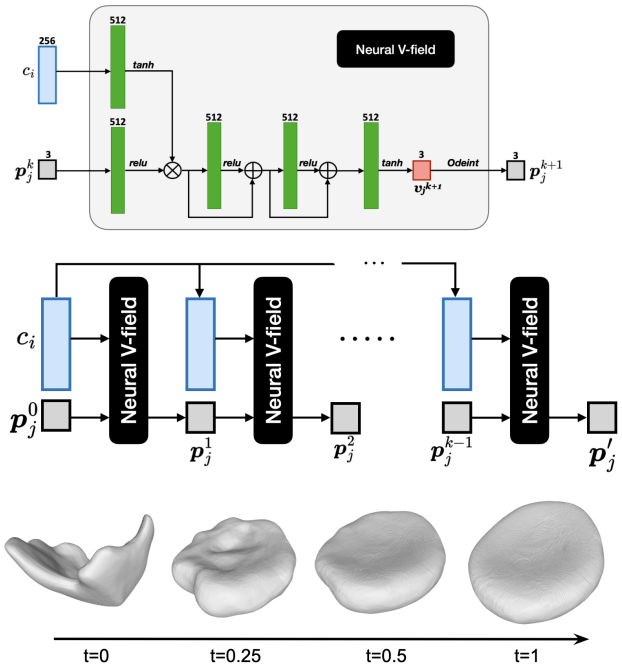

Figure 3. Neural ODE–based deformation module. Top: a velocity field is predicted by a residual MLP from the current point position and latent code c_i, and integrated with an ODE solver to produce the updated position. Middle: The full deformation Φ_i is obtained by concatenating four NODE blocks, where the output of block k serves as the input to block $k + 1$. Bottom: Visualization of the progressive deformation steps, illustrating the smooth, monotonic trajectory that warps each placenta instance into the learned template space.

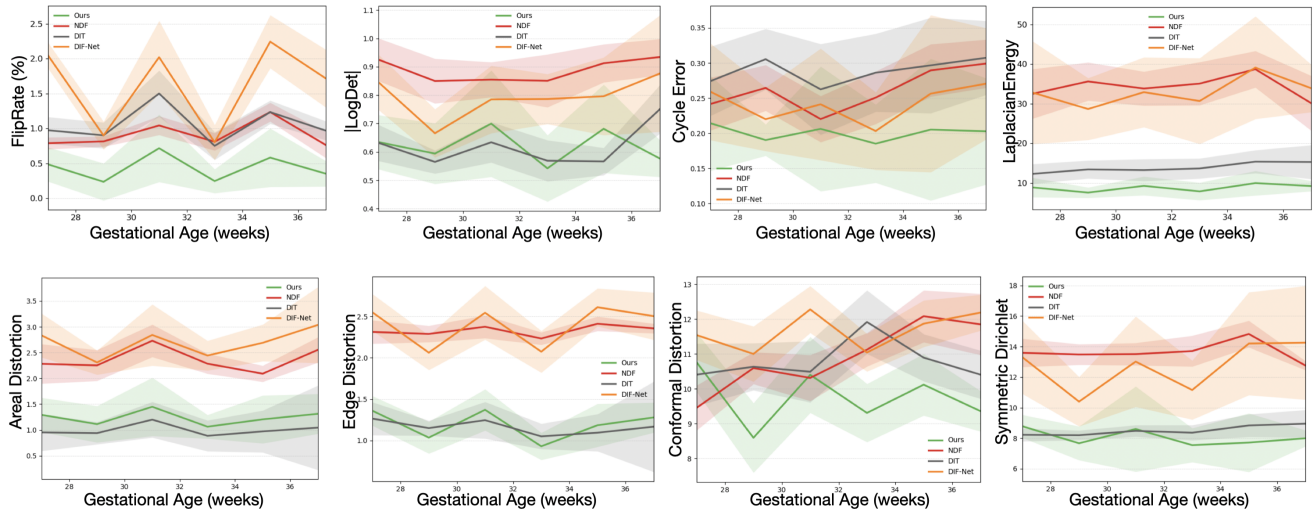

Figure 4. Distribution of diffeomorphism and distortion metrics across gestational age.

umetric deformations throughout late pregnancy (Fig. 4).

Distortion Analysis. We evaluate geometric distortion using volumetric, areal, and edge stretch ratios, together with symmetric Dirichlet energy [5, 6] and conformal distortion [2]. All distortion measures are dimensionless and attain a value of 1 for perfect volume-, area-, or length-preserving mappings (the symmetric Dirichlet energy attains its isometric minimum at 6). Across all metrics, our method exhibits lower distortion than baseline methods (Fig. 4). These improvements result in deformations that are better-conditioned and more regularized, reducing directional distortion throughout the volume and yielding smoother, more volume-preserving behavior.

References

[1] Kunal Gupta. *Neural mesh flow: 3d manifold mesh generation via diffeomorphic flows*. University of California, San Diego, 2020. 1

[2] Bruno Lévy, Sylvain Petitjean, Nicolas Ray, and Jérome Maillot. Least squares conformal maps for automatic texture atlas generation. In *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pages 193–202. 2023. 3

[3] William E Lorensen and Harvey E Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *Seminal graphics: pioneering efforts that shaped the field*, pages 347–353. 1998. 1

[4] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. Deepsdf: Learning continuous signed distance functions for shape representation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 165–174, 2019. 1

[5] John Schreiner, Arul Asirvatham, Emil Praun, and Hugues Hoppe. Inter-surface mapping. In *ACM SIGGRAPH 2004 Papers*, pages 870–877. 2004. 3

[6] Jason Smith and Scott Schaefer. Bijective parameterization with free boundaries. *ACM Transactions on Graphics (TOG)*, 34(4):1–9, 2015. 3

[7] Shanlin Sun, Kun Han, Deying Kong, Hao Tang, Xiangyi Yan, and Xiaohui Xie. Topology-preserving shape reconstruction and registration via neural diffeomorphic flow. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 20845–20855, 2022. 1

[8] Zerong Zheng, Tao Yu, Qionghai Dai, and Yebin Liu. Deep implicit templates for 3d shape representation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1429–1439, 2021. 1