Weakly Supervised Learning of Cortical Surface Reconstruction from Segmentations

Qiang Ma[1][0000−0003−0791−1731], Liu Li[1][0000−0003−2376−8162], Emma C. Robinson[2][0000−0002−7886−3426], Bernhard Kainz[1,2,3][0000−0002−7813−5023], and Daniel Rueckert[1,4][0000−0002−5683−5889]

[1] Department of Computing, Imperial College London, UK
q.ma20@imperial.ac.uk
[2] King's College London, UK
[3] FAU Erlangen–Nürnberg, Germany
[4] Klinikum rechts der Isar, Technical University of Munich, Germany

Abstract. Existing learning-based cortical surface reconstruction approaches heavily rely on the supervision of pseudo ground truth (pGT) cortical surfaces for training. Such pGT surfaces are generated by traditional neuroimage processing pipelines, which are time consuming and difficult to generalize well to low-resolution brain MRI, *e.g.*, from fetuses and neonates. In this work, we present CoSeg, a learning-based cortical surface reconstruction framework weakly supervised by brain segmentations without the need for pGT surfaces. CoSeg introduces temporal attention networks to learn time-varying velocity fields from brain MRI for diffeomorphic surface deformations, which fit an initial surface to target cortical surfaces within only 0.11 seconds for each brain hemisphere. A weakly supervised loss is designed to reconstruct pial surfaces by inflating the white surface along the normal direction towards the boundary of the cortical gray matter segmentation. This alleviates partial volume effects and encourages the pial surface to deform into deep and challenging cortical sulci. We evaluate CoSeg on 1,113 adult brain MRI at 1mm and 2mm resolution. CoSeg achieves superior geometric and morphological accuracy compared to existing learning-based approaches. We also verify that CoSeg can extract high-quality cortical surfaces from fetal brain MRI on which traditional pipelines fail to produce acceptable results.

Keywords: Cortical Surface Reconstruction · Weak Supervision · MRI

1 Introduction

Cortical surfaces, *i.e.*, the inner/white and outer/pial surfaces of the cerebral cortex, play a crucial role in visualizing the anatomical structure as well as quantitative characterizing the morphology of the cortex. Traditional neuroimage processing pipelines [9,11,12,21] achieved great success for cortical surface reconstruction from adult or neonatal brain MRI. However, these pipelines normally incorporate multiple processing steps and require more than 6 hours for a single subject. The limited accuracy of prior processing steps will also cause

subsequent corruptions in cortical surfaces. Moreover, traditional pipelines contain a large amount of carefully tuned parameters, which make them difficult to generalize across data domains in different age groups or acquisition protocols.

To boost the accuracy and efficiency of cortical surface reconstruction, recent studies [8,13,14,32] leverage deep learning-based approaches to predict implicit surface representations, *i.e.*, segmentations and signed distance functions (SDF). Explicit 3D meshes are extracted from the implicit surfaces by Marching Cubes (MC) algorithm [17]. Topology correction algorithms [1,27] are used to detect and repair topological errors such that the extracted surfaces are homeomorphic to a 2-sphere. However, this is a time-consuming process based on iterative refinement, which is hard to accelerate efficiently through engineering efforts. In addition, although existing approaches are capable of producing accurate segmentations [2,14,24,29,32] or SDFs [8,13], these volumetric representations intrinsically face the challenges of partial volume effects, leading to difficulties for the MC algorithm in accurately capturing the cortical folding of extracted surfaces. This issue is especially pronounced within the cortical sulci, in particular for low-resolution brain MRI such as fetal and neonatal data.

Instead of learning implicit surfaces, latest works [3,4,6,15,16,18,19,20,25,34] focus on learning explicit surface deformations end-to-end from brain MRI. Such approaches only require a few seconds to deform an initial mesh to target cortical surfaces. However, one major limitation is that they heavily rely on the supervision of pseudo ground truth (pGT) cortical surfaces generated by traditional pipelines [9,11,12,21] such as FreeSurfer [11]. Their long processing time makes it expensive to collect a large dataset for training. Also, errors in pGT surfaces inevitably form an upper bound for the achievable accuracy. Moreover, traditional pipelines may fail to extract pGT surfaces from fetal or clinical MRI sequences, which are inherently limited in scan time and thus lower image resolution.

Contributions: We present CoSeg, a weakly supervised cortical surface reconstruction framework without the need for pGT surfaces during the training phase. CoSeg learns cortical surfaces explicitly with the weak supervision of cortical ribbon segmentations, which are more accessible than pGT cortical surfaces particularly for low-resolution images [2,29]. We introduce temporal attention networks (TA-Net) [18] to learn diffeomorphic surface deformations from brain MRI. To address partial volume effects on the pial surfaces, we propose a novel weakly supervised loss consisting of a boundary loss, which inflates the white surface towards the boundary of the cortical gray matter (cGM) segmentation, as well as an inflation loss, which constraints the inflation to follow the normal direction. We evaluate CoSeg on the HCP young adult dataset [31] and provide qualitative evaluation results for the dHCP fetal dataset [22]. The code for CoSeg are released publicly at https://github.com/m-qiang/CoSeg.

2   Method

Temporal Attention Networks. Given an initial surface $\mathcal{S}_0 \subset \mathbb{R}^3$ with points $x_0 \in \mathcal{S}_0$, a diffeomorphic surface deformation can be defined via an ordinary
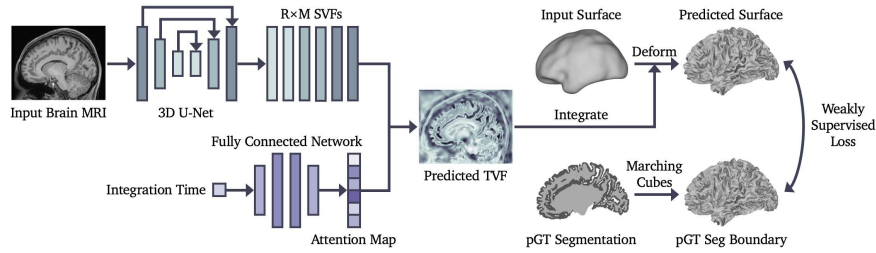
Fig. 1. The architecture of the TA-Net. TA-Net predicts a TVF by learning a weighted sum of multiple SVFs. The TVF is integrated to produce diffeomorphic surface deformations which fit an initial input surface to the predicted cortical surface.

differential equation (ODE):

$$\dot{x}(t) = u(x(t), t), \; x(0) = x_0, \; t \in [0, T], \tag{1}$$

where $x(t)$ is the deformed surface point, and $u(x, t) \in \mathbb{R}^3$ is a learnable time-varying velocity field (TVF). By integrating the flow ODE (1), the initial surface $\mathcal{S}_0$ is deformed to a target surface $\mathcal{S}_T$ with the points $x_T = x(T)$.

Since our CoSeg framework is model agnostic, we introduce a temporal attention network (TA-Net) based on CoTAN [18] to learn the TVF $u(x, t)$. While CoTAN is conditioned on the ages of the subjects, TA-Net is only conditioned on the integration time t such that it can be extended to adult data. As shown in Fig. 2, given an input brain MRI image, TA-Net uses a 3D U-Net [23] to learn $M = 2$ stationary velocity fields (SVF) for each of $R = 3$ resolution levels. Thus, there are total $R \times M$ SVFs denoted by $\mathbf{u}(x) \in \mathbb{R}^{RM \times 3}$. Given an input integration time $t \in [0, T]$, a channel-wise time-varying attention map $\mathbf{p}(t) \in \mathbb{R}^{RM}$, which satisfies $\sum_{i=1}^{RM} \mathbf{p}_i(t) = 1$, is predicted by a fully connected network to measure the importance of all $R \times M$ SVFs. By computing the weighted sum over all SVFs, the TVF is represented as $u(x, t) = \mathbf{u}(x)^\top \mathbf{p}(t)$.

We discretize a surface $\mathcal{S}$ by a 3D mesh $\mathcal{M} = (\mathcal{V}, \mathcal{F}, \mathcal{E})$, where $\mathcal{V}, \mathcal{F}, \mathcal{E}$ are the sets of vertices, faces, and edges respectively. Then, we integrate the ODE (1) with the forward Euler method. For each integration step $k = 0, ..., K - 1$, the vertices can be updated by $v_{k+1} = v_k + hu(v_k, hk)$, where $h = T/K$ is the step size and $v_0 \in \mathcal{V}_0$ is the vertex of an input surface mesh $\mathcal{M}_0$. By integration, the TA-Net deforms the input mesh $\mathcal{M}_0$ to a predicted cortical surface mesh $\hat{\mathcal{M}}$ with vertices $\hat{v} = v_K$. The TA-Net can be trained by minimizing a weakly supervised loss between the predicted mesh $\hat{\mathcal{M}}$ and the pGT segmentation boundary.

Initial Surface Generation. We affinely align all brain MRI and pGT cortical ribbon segmentations to a standard space such as MNI-152. A fixed initial surface for all subjects is extracted from a template ribbon segmentation, which is obtained from either an atlas or the group average of all pGT segmentations in the training set. Inspired by [19,25], we use a distance transform algorithm [5] to convert the template white matter (WM) segmentation to a SDF, where the

interior/exterior of the segmentation has negative/positive value. The SDF is smoothed by a Gaussian filter with standard deviation σ=6.0. An initial mesh $\mathcal{M}_0$ is extracted by MC [17] from the SDF at level l=1.5. Such an initial mesh has genus-0 topology since the implicit surface has been sufficiently smoothed and inflated. We iteratively apply Laplacian smoothing and isotropic explicit remeshing [7] to the initial mesh $\mathcal{M}_0$ until it has the desired number of vertices.

White Surface Reconstruction. CoSeg learns cortical surfaces weakly supervised by cortical ribbon segmentations, which contain the label maps of WM and cGM. The WM label includes subcortical structures such as deep gray matter and ventricle. The pGT segmentations can be obtained from either advanced learning-based approaches [2,9,14,24,29] or manual annotations.

To learn the white surface, we minimize the commonly used bidirectional Chamfer (bi-Chamfer) distance $\mathcal{L}_{\text{chamfer}}$ [3,10,16,18,19,25,33] between the predicted white surface mesh $\hat{\mathcal{M}}$ and the boundary of the pGT WM segmentation. The boundary is represented as a 3D surface mesh $\mathcal{M}_*$ extracted from the WM segmentation by MC [17]. Taubin smoothing [28] is applied to alleviate the grid artifacts of the extracted mesh. Then, the bi-Chamfer loss [10] is defined as:

$$\mathcal{L}_{\text{chamfer}}(\hat{\mathcal{M}}, \mathcal{M}_*) = \frac{1}{|\hat{\mathcal{V}}|} \sum_{\hat{v} \in \hat{\mathcal{V}}} \min_{v_* \in \mathcal{V}_*} \|\hat{v} - v_*\|^2 + \frac{1}{|\mathcal{V}_*|} \sum_{v_* \in \mathcal{V}_*} \min_{\hat{v} \in \hat{\mathcal{V}}} \|v_* - \hat{v}\|^2,$$
(2)

where $\hat{v}$ and v_* are the vertices of the predicted mesh $\hat{\mathcal{M}}$ and the pGT WM segmentation boundary $\mathcal{M}_*$ respectively. We add the edge length loss $\mathcal{L}_{\text{edge}}$ and normal consistency loss $\mathcal{L}_{\text{nc}}$ [3,33] to enforce surface smoothness. The final loss is defined as $\mathcal{L}_{\text{white}} = \mathcal{L}_{\text{chamfer}} + w_{\text{edge}} \mathcal{L}_{\text{edge}} + w_{\text{nc}} \mathcal{L}_{\text{nc}}$ with weights w_{edge} and w_{nc}.

Pial Surface Reconstruction. Since the cGM segmentations are strongly affected by partial volume effects, it only provides weak supervision especially for low-resolution data. Fig. 2-a shows that the cGM segmentation fails to identify the voxels in the deep cortical sulci. Different from white surface reconstruction, the pial surfaces learned through the bi-Chamfer loss may fail to capture cortical folding and thus cause inaccurate estimation of morphological features.

Inspired by deformation-based approaches [21,26], we propose a novel weakly supervised loss to tackle partial volume effects on the pial surfaces. The predicted white surface is used as the input mesh $\mathcal{M}_0$. First, we introduce a *boundary loss*:

$$\mathcal{L}_{\text{boundary}}(\hat{\mathcal{M}}, \mathcal{M}_*) = \frac{2}{|\mathcal{V}_*|} \sum_{v_* \in \mathcal{V}_*} \min_{\hat{v} \in \hat{\mathcal{V}}} \|v_* - \hat{v}\|^2. $$
(3)

The boundary loss is equivalent to a single-directional Chamfer loss that computes the shortest distance from the boundary $\mathcal{M}_*$ of the pGT cGM segmentation to the predicted pial surface $\hat{\mathcal{M}}$. As shown in Fig. 2-c,d, the proposed boundary loss will extend the input white surface towards the cGM boundary without affecting the sulcal regions, whereas the bi-Chamfer distance will deform
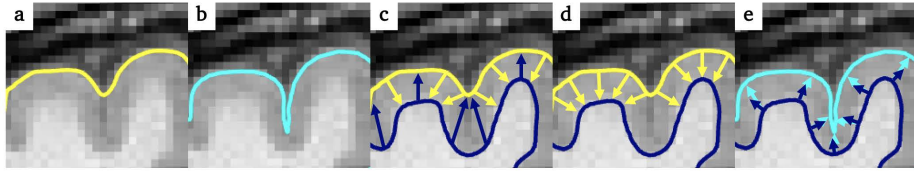
Fig. 2. T1w brain MRI from the HCP dataset [31]. (a) The cGM segmentation boundary (yellow); (b) The expected pial surface (cyan); (c) Bi-Chamfer distance between the cGM segmentation boundary and input white surface (blue). (d) The boundary loss, *i.e.*, single-directional Chamfer distance. (e) The inflation loss between the vertex displacement (cyan) and the normal vector (blue) of the input white surface.

the surface outwards from the deep sulci. Next, an *inflation loss* is defined as:

$$\mathcal{L}_{\text{inflation}}(\hat{\mathcal{M}}, \mathcal{M}_0) = 1 - \frac{1}{|\hat{\mathcal{V}}|} \sum_{i=1}^{|\hat{\mathcal{V}}|} \frac{\hat{v}^i - v_0^i}{\|\hat{v}^i - v_0^i\| + \epsilon} \cdot n(v_0^i), \tag{4}$$

where $\epsilon=10^{-12}$, v_0^i is the i-th vertex of the input white surface $\mathcal{M}_0$, and $n(v_0^i)$ is the normal vector with $\|n(v_0^i)\|=1$. The inflation loss measures cosine similarity between the vertex displacement and the normal vectors (Fig. 2-e). This ensures that the inflation of input white surface adheres to its normal direction. The final weakly supervised loss is defined as $\mathcal{L}_{\text{weak}} = \mathcal{L}_{\text{boundary}} + w_{\text{inflation}}\mathcal{L}_{\text{inflation}}$.

The weakly supervised loss encourages the input white surface to inflate along the normal until it fits the cGM boundary or two gyri touch each other. Such inflation can alleviate the partial volume effects and allows the pial surfaces to deform into the deep cortical sulci. During the inflation, mesh self-intersections can be prevented by diffeomorphic surface deformations without additional collision detection. Considering the smoothness terms, the final loss for pial surface reconstruction is defined as $\mathcal{L}_{\text{pial}} = \mathcal{L}_{\text{weak}} + w_{\text{edge}}\mathcal{L}_{\text{edge}} + w_{\text{nc}}\mathcal{L}_{\text{nc}}$.

At the beginning of the training, we have $\hat{\mathcal{M}} = \mathcal{M}_0$ since the input white surface has not been deformed. The gradient of the inflation loss $\mathcal{L}_{\text{inflation}}$ with respect to the i-th vertex can be computed as $\nabla_i \mathcal{L}_{\text{inflation}}|_{\hat{\mathcal{M}}=\mathcal{M}_0} = -n(v_0^i)/(\epsilon|\hat{\mathcal{V}}|)$. However, this leads to exploding gradients as ϵ is very small. To address this issue, we pre-train the TA-Net for a few epochs by replacing $\mathcal{L}_{\text{weak}}$ with a MSE loss $\mathcal{L}_{\text{mse}}(\hat{\mathcal{M}}, \mathcal{M}_N) = \sum_{i=1}^{|\hat{\mathcal{V}}|} \|\hat{v}^i - v_N^i\|^2/|\hat{\mathcal{V}}|$, where $\mathcal{M}_N$ is an inflated white surface with vertices $v_N^i = v_{N-1}^i + \gamma n(v_{N-1}^i)$ for $N=1, ..., 10$ and $\gamma=0.1$. The predicted surface learns to extend outwards during pre-training while avoiding exploding gradients. After pre-training, we resume to use $\mathcal{L}_{\text{weak}}$ as the reconstruction loss.

3   Experiments

HCP Young Adult Dataset. We evaluate CoSeg on the publicly available HCP young adult dataset [31] with 1,113 T1w brain MRI, cortical ribbon segmentations and cortical surfaces extracted by the HCP pipeline [12] with quality

Table 1. Comparison with existing learning-based approaches in terms of the types of surface representation, required supervision and primary loss functions.

Table 2. Comparison with implicit approaches on the HCP dataset at 1mm/2mm resolutions. The ASSD (mm) and HD90 (mm) are reported.

Method	Type	pGT	Loss function
3D U-Net [23]	Implicit	Seg	Cross Entropy
DeepCSR [8]	Implicit	SDF	L1 Loss
Vox2Cortex [3]	Explicit	Surface	Bi-Chamfer
CortexODE [19]	Explicit	Surface	Bi-Chamfer
CFPP [25]	Explicit	Surface	Bi-Chamfer
CoTAN [18]	Explicit	Surface	Bi-Chamfer
CoSeg (Ours)	Explicit	Seg	Weakly Supervised

|	Method	White		Pial	
		ASSD↓	HD90↓	ASSD↓	HD90↓
1mm	3D U-Net	0.180	**0.373**	0.432	1.257
	DeepCSR	0.213	0.468	0.436	1.456
	CoSeg	**0.177**	0.389	**0.257**	**0.553**
2mm	3D U-Net	0.385	0.831	0.833	3.125
	DeepCSR	0.404	1.019	0.904	4.007
	CoSeg	**0.330**	**0.756**	**0.361**	**0.767**
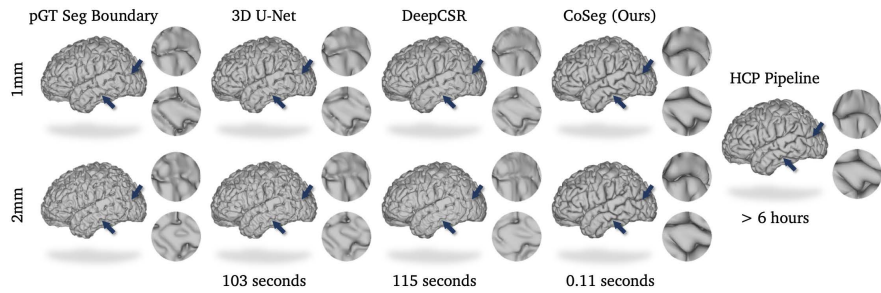
Fig. 3. Pial surfaces predicted by implicit approaches on the HCP dataset. The runtime required for each brain hemisphere (including topology correction) is reported.

control. The segmentations are used as the pGT for training, and the pGT cortical surfaces are only used for evaluation. The dataset is split by the ratio of 60/10/30% for training/validation/testing. All T1w images and segmentations have been affinely aligned to the MNI-152 space. We compute the average segmentations of all training subjects as a template, from which we extract an initial surface with 160k vertices. Both MRI and segmentations are resampled to 1mm/2mm isotropic resolutions and clipped to the sizes of $112\times224\times176$ and $56\times112\times88$ for each brain hemisphere. Without loss of generality, we only consider the left hemisphere in the experiments. We train TA-Nets for 200 epochs by Adam optimizer with learning rate 0.0001. For pial surfaces, we first pre-train the TA-Net using the MSE loss $\mathcal{L}_{\mathrm{mse}}$ for 20 epochs, and then optimize the weakly supervised loss $\mathcal{L}_{\mathrm{weak}}$ for 180 epochs. We use w_{edge}=0.5 and w_{nc}=5.0 for smoothness terms. The integration time and step size are set to T=1 and h=0.02 with K=50 steps. We compare CoSeg with both implicit and explicit learning-based cortical surface reconstruction approaches as shown in Table 1. All experiments are performed on a NVIDIA RTX3080 GPU with 10GB memory.

— *Comparison with Implicit Approaches:* We compare CoSeg with implicit deep learning-based approaches for 1mm and 2mm resolutions. For CoSeg, we train two TA-Nets for white and pial surfaces supervised by the pGT segmentation boundary. The weights $w_{\mathrm{inflation}}$ are set to 2.0/5.0 for 1mm/2mm resolutions. Since the main backbone of the TA-Net is a 3D U-Net [23], we train U-Nets

Table 3. Comparison with explicit learning-based approaches and ablation studies of CoSeg on the HCP dataset at 1mm/2mm resolutions. The ASSD (mm), HD90 (mm), errors of cortical thickness e_{thick} (mm) and sulcal depth e_{sulc} (mm) are reported. The best results are in bold and the second best results are underlined.

Method	Resolution (1mm)				Resolution (2mm)			
	ASSD$\downarrow$	HD90$\downarrow$	$e_{\mathrm{thick}}\downarrow$	$e_{\mathrm{sulc}}\downarrow$	ASSD$\downarrow$	HD90$\downarrow$	$e_{\mathrm{thick}}\downarrow$	$e_{\mathrm{sulc}}\downarrow$
Vox2Cortex [3]	0.292	0.641	0.511	1.238	0.542	2.026	0.731	1.432
CortexODE [19]	0.352	0.796	0.448	0.672	0.610	1.787	0.742	1.303
CFPP [25]	0.274	0.671	0.496	1.417	0.578	2.385	0.872	1.647
CoTAN [18]	0.293	0.769	0.534	1.438	0.538	1.966	0.781	1.781
CoSeg (w_{inflate}=0.0)	<u>0.253</u>	<u>0.564</u>	0.379	1.389	0.383	0.980	0.551	1.580
CoSeg (w_{inflate}=2.0)	**0.252**	**0.539**	**0.298**	<u>0.530</u>	**0.347**	**0.769**	0.424	<u>0.686</u>
CoSeg (w_{inflate}=5.0)	0.273	0.627	<u>0.320</u>	**0.435**	<u>0.351</u>	<u>0.777</u>	**0.386**	**0.397**
CoSeg (w_{inflate}=10.0)	0.298	0.736	0.362	1.120	0.370	0.881	<u>0.421</u>	0.797
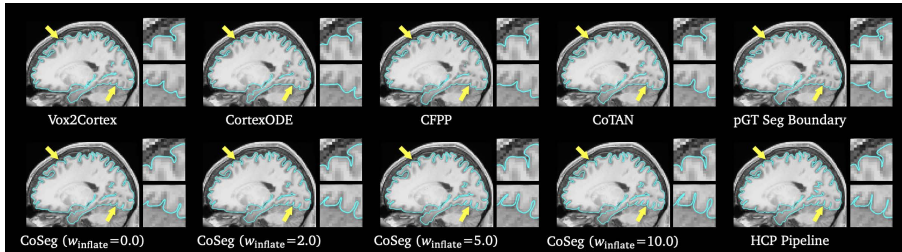
Fig. 4. Pial surfaces (2mm) predicted by explicit approaches on the HCP dataset.

to learn segmentations for fair comparison. We also compare to DeepCSR [8], of which the pGT SDFs are created by the distance transform [5] of pGT segmentations. The surface meshes ($|V|\approx$150k) are extracted by MC followed by topology correction [1] and Taubin smoothing [28]. The performance is evaluated by the average symmetric surface distance (ASSD) and 90th-percentile Hausdorff distance (HD90) [3,8,19] between the predicted and pGT cortical surfaces.

Table 2 reports that CoSeg achieves superior geometric accuracy except for white surfaces at 1mm. Fig. 3 shows that CoSeg is able to deform the pial surfaces into deep sulci, while the implicit approaches fail to capture the cortical sulci due to intrinsic partial volume effects of volumetric data, leading to large geometric errors for pial surfaces. CoSeg only requires 0.11s of runtime for each brain hemisphere. This is orders of magnitude faster than the HCP pipeline, U-Net and DeepCSR as reported in Fig. 3. CoSeg only produces a negligible number of self-intersecting faces, *i.e.*, 0.06/3.65 out of 330k faces on average for white/pial surface respectively.

— *Comparison with Explicit Approaches:* We compare CoSeg with explicit deep learning-based approaches including Vox2Cortex [3], CortexODE [19], CorticalFlow++ (CFPP) [25] and CoTAN [18]. For CoTAN we use the same architecture as the TA-Net for adult data. To show the advantage of proposed weakly supervised loss, we train all approaches to learn pial surfaces from the pGT segmentation boundary at 1mm/2mm resolutions. The pGT white surfaces are used
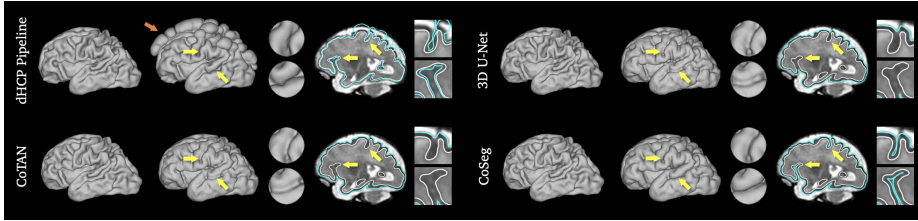
Fig. 5. Cortical surfaces of a 32.4-week subject in the dHCP fetal dataset.

as the input for training, such that the predicted and pGT pial surfaces have the same mesh connectivity. This allows vertex-to-vertex comparison of cortical morphological features [12]. We compute the average L1 errors of the cortical thickness and the sulcal depth of midthickness surfaces over all vertices between the predicted and pGT cortical surfaces. The geometric accuracy is measured by ASSD and HD90. We also conduct ablation studies to explore the influence of different weights $w_{\text{inflation}}$ for the inflation loss $\mathcal{L}_{\text{inflation}}$ on CoSeg.

As shown in Table 3 and Fig. 4, CoSeg achieves superior geometric and morphological accuracy. Existing explicit approaches [3,18,19,25], which use the bi-Chamfer loss (see Table 1), overfit to the pGT segmentation boundary and thus fail to fit the deep cortical sulci especially for low resolution samples. This results in inaccurate morphological features, *i.e.*, greater cortical thickness and shallower sulcal depth (see Appendix). Our ablation study shows that the inflation loss effectively reduces morphological errors. However, excessive weight for the inflation loss also affects the performance, since the vertex displacement of the pial surface may not strictly follow the normal direction of the white surface.

dHCP Fetal Dataset. We evaluate CoSeg on the dHCP fetal dataset [22] with 241 T2w fetal brain MRI scanned at the ages of 22–38 weeks. The dataset is split by the ratio of 60/10/30%. The voxel size of the T2w images is 0.5mm^3. However, this is a still relatively low resolution as the fetal brains are very small. The pGT tissue segmentations are created by the BOUNTI pipeline [29], a fully validated learning-based fetal brain segmentation approach. The pGT cortical ribbon segmentations are created by merging the tissue labels. Both T2w images and segmentations are affinely aligned to a 36-week fetal brain MRI atlas [30] and clipped to the size of $112\times224\times176$ for each hemisphere. An initial surface is generated based on the segmentation atlas [30]. For CoSeg we set $w_{\text{inflation}}$=5.0.

We compare CoSeg with the dHCP neonatal pipeline [21], CoTAN [18] and 3D U-Net supervised by pGT segmentations. Since the Dice score cannot reflect the geometric accuracy, we provide qualitative comparisons in Fig. 5. It shows that CoSeg predicts high-quality surfaces, while the dHCP neonatal pipeline fails to generalize well on fetal subjects. The pial surfaces predicted by U-Net and CoTAN are severely affected by the partial volume effects of the cGM. Compared to the dHCP neonatal pipeline which requires 6.5 hours of runtime, CoSeg only needs less than 0.2 seconds for each brain hemisphere.

4   Conclusion

In this work, we propose CoSeg for diffeomorphic cortical surface reconstruction weakly supervised by cortical ribbon segmentations. A novel weakly supervised loss is designed to address partial volume effects and deform the pial surfaces into deep cortical sulci. CoSeg shows remarkable results on both HCP young adult and dHCP fetal dataset without the need for pGT cortical surfaces. One limitation is that CoSeg guides vertex displacement strictly along the normal of the input white surface, with potential to influence anatomical fidelity of pial surfaces. The efficacy of CoSeg is also affected by the quality of pGT segmentations. Looking ahead, we plan to explore semi-supervised learning, utilizing both labeled segmentations and unlabeled MRI intensity to refine the performance.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bazin, P.L., Pham, D.L.: Topology correction using fast marching methods and its application to brain segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 484–491. Springer (2005)
2. Billot, B., Greve, D.N., et al.: SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining. Medical image analysis **86**, 102789 (2023)
3. Bongratz, F., Rickmann, A.M., Pölsterl, S., Wachinger, C.: Vox2Cortex: fast explicit reconstruction of cortical surfaces from 3D MRI scans with geometric deep neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20773–20783 (2022)
4. Bongratz, F., Rickmann, A.M., Wachinger, C.: Neural deformation fields for template-based reconstruction of cortical surfaces from MRI. Medical Image Analysis p. 103093 (2024)
5. Breu, H., Gil, J., Kirkpatrick, D., Werman, M.: Linear time Euclidean distance transform algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence **17**(5), 529–533 (1995)
6. Chen, X., Zhao, J., Liu, S., Ahmad, S., Yap, P.T.: SurfFlow: A flow-based approach for rapid and accurate cortical surface reconstruction from infant brain MRI. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 380–388. Springer (2023)

7. Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., Ranzuglia, G., et al.: MeshLab: an open-source mesh processing tool. In: Eurographics Italian chapter conference. vol. 2008, pp. 129–136. Salerno, Italy (2008)

8. Cruz, R.S., Lebrat, L., et al.: DeepCSR: A 3D deep learning approach for cortical surface reconstruction. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 806–815 (2021)

9. Dai, Y., Shi, F., Wang, L., Wu, G., Shen, D.: iBEAT: a toolbox for infant brain magnetic resonance image processing. Neuroinformatics **11**, 211–225 (2013)

10. Fan, H., Su, H., Guibas, L.J.: A point set generation network for 3D object reconstruction from a single image. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 605–613 (2017)

11. Fischl, B.: FreeSurfer. Neuroimage **62**(2), 774–781 (2012)

12. Glasser, M.F., Sotiropoulos, S.N., et al.: The minimal preprocessing pipelines for the human connectome project. Neuroimage **80**, 105–124 (2013)

13. Gopinath, K., Greve, D.N., et al.: Cortical analysis of heterogeneous clinical brain MRI scans for large-scale neuroimaging studies. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer (2023)

14. Henschel, L., Conjeti, S., et al.: FastSurfer - a fast and accurate deep learning based neuroimaging pipeline. NeuroImage **219**, 117012 (2020)

15. Hoopes, A., Iglesias, J.E., et al.: TopoFit: rapid reconstruction of topologically-correct cortical surfaces. Proceedings of machine learning research **172**, 508 (2022)

16. Lebrat, L., Santa Cruz, R., et al.: CorticalFlow: a diffeomorphic mesh transformer network for cortical surface reconstruction. Advances in Neural Information Processing Systems **34**, 29491–29505 (2021)

17. Lorensen, W.E., Cline, H.E.: Marching Cubes: A high resolution 3D surface construction algorithm. In: Seminal graphics: pioneering efforts that shaped the field, pp. 347–353 (1998)

18. Ma, Q., Li, L., Kyriakopoulou, V., Hajnal, J.V., Robinson, E.C., Kainz, B., Rueckert, D.: Conditional temporal attention networks for neonatal cortical surface reconstruction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 312–322. Springer (2023)

19. Ma, Q., Li, L., Robinson, E.C., Kainz, B., Rueckert, D., Alansary, A.: CortexODE: Learning cortical surface reconstruction by neural ODEs. IEEE Transactions on Medical Imaging **42**(2), 430–443 (2022)

20. Ma, Q., Robinson, E.C., Kainz, B., Rueckert, D., Alansary, A.: PialNN: a fast deep learning framework for cortical pial surface reconstruction. In: Machine Learning in Clinical Neuroimaging. pp. 73–81. Springer (2021)

21. Makropoulos, A., Robinson, E.C., Schuh, A., et al.: The developing human connectome project: A minimal processing pipeline for neonatal cortical surface reconstruction. Neuroimage **173**, 88–112 (2018)

22. Price, A.N., Cordero-Grande, L., et al.: The developing human connectome project (dHCP): fetal acquisition protocol. In: Proceedings of the annual meeting of the International Society of Magnetic Resonance in Medicine (ISMRM). vol. 244 (2019)

23. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241. Springer (2015)

24. Roy, A.G., Conjeti, S., Navab, N., Wachinger, C., Initiative, A.D.N., et al.: QuickNAT: A fully convolutional network for quick and accurate segmentation of neuroanatomy. NeuroImage **186**, 713–727 (2019)

25. Santa Cruz, R., Lebrat, L., Fu, D., Bourgeat, P., Fripp, J., Fookes, C., Salvado, O.: CorticalFlow++: Boosting cortical surface reconstruction accuracy, regularity, and interoperability. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 496–505. Springer (2022)

26. Schuh, A., Makropoulos, A., Wright, R., Robinson, E.C., et al.: A deformable model for the reconstruction of the neonatal cortex. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). pp. 800–803. IEEE (2017)

27. Ségonne, F., Pacheco, J., Fischl, B.: Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. IEEE transactions on medical imaging **26**(4), 518–529 (2007)

28. Taubin, G.: Curve and surface smoothing without shrinkage. In: Proceedings of IEEE international conference on computer vision. pp. 852–857. IEEE (1995)

29. Uus, A.U., Kyriakopoulou, V., et al.: BOUNTI: Brain volumetry and automated parcellation for 3D fetal MRI. Elife (2023)

30. Uus, A.U., Kyriakopoulou, V., et al.: Multi-channel spatio-temporal MRI atlas of the normal fetal brain development from the developing human connectome project. G-Node (2023). https://doi.org/https://doi.org/10.12751/g-node.ysgsy1

31. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.M.H., et al.: The WU-Minn human connectome project: an overview. Neuroimage **80**, 62–79 (2013)

32. Wang, L., Wu, Z., Chen, L., Sun, Y., Lin, W., Li, G.: iBEAT V2.0: a multisite-applicable, deep learning-based pipeline for infant cerebral cortical surface reconstruction. Nature protocols **18**(5), 1488–1509 (2023)

33. Wickramasinghe, U., Remelli, E., Knott, G., Fua, P.: Voxel2Mesh: 3D mesh model generation from volumetric data. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 299–308. Springer (2020)

34. Zheng, H., Li, H., Fan, Y.: Coupled reconstruction of cortical surfaces by diffeomorphic mesh deformation. Advances in Neural Information Processing Systems **36** (2024)