Poster No 1629

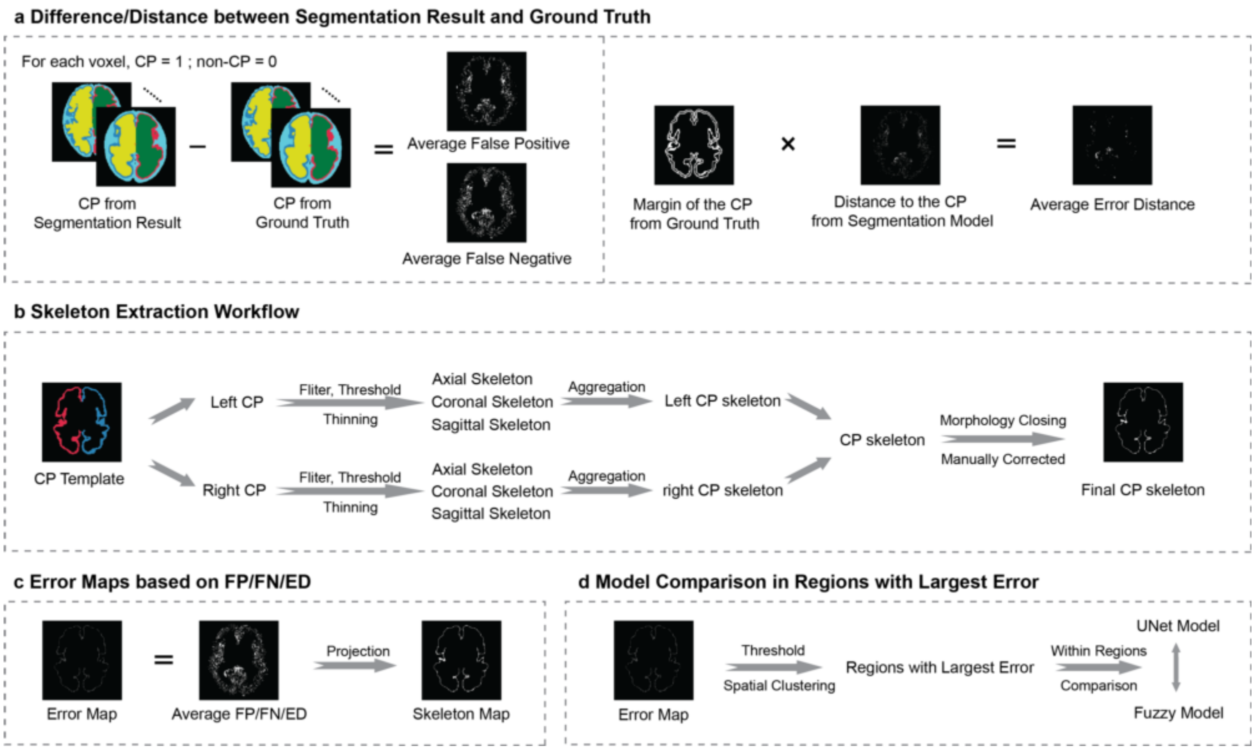
Regional Analysis of Error Patterns from Automatic Cortical Plate Segmentation in Fetal Brain MRI

Bowen Qiu[1,2], Andrea Gondova[1,3], Sungmin You[1,3], Milton Candela-Leal[1], Hanuman Verma[1], Seungyoon Jeong[1,3], P. Ellen Grant[1,3,4], Kiho Im[1,3,5]

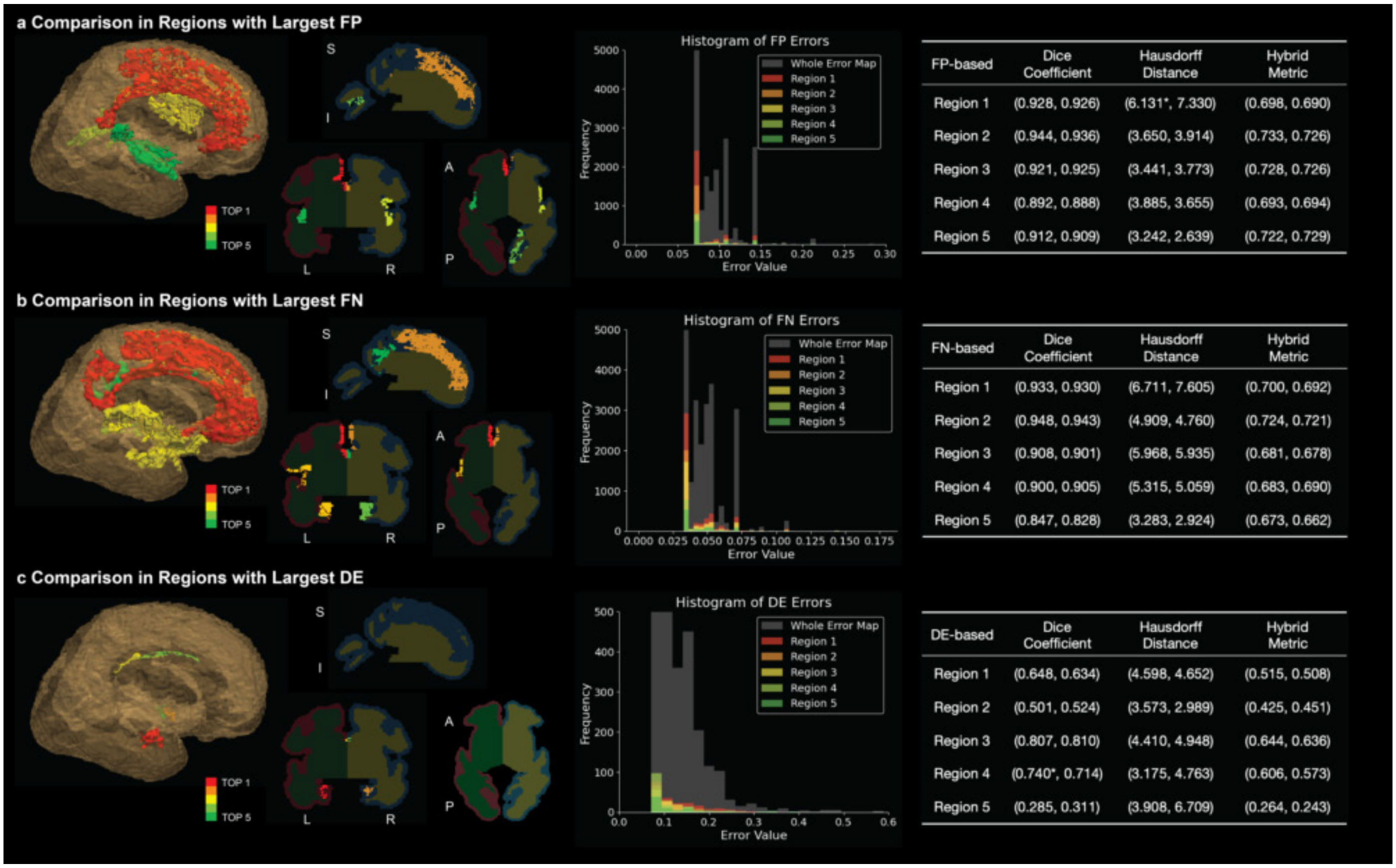
[1]Fetal Neonatal Neuroimaging and Developmental Science Center, Harvard Medical School, Boston, MA, USA, [2]Key Laboratory for Biomedical Engineering of Ministry of Education, Department of Biomedical Engineering, College of Biomedical Engineering & Instrument Science, Zhejiang University, Hangzhou, Zhejiang, China, [3]Division of Newborn Medicine, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA, [4]Department of Radiology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA, [5]Department of Pediatrics, Harvard Medical School, Boston, MA, USA

Introduction: Fetal brain segmentation, with a specific focus on the cortical plate (CP), is essential for the early detection of atypical brain development. Recent advances, deep learning approaches such as convolutional neural networks (CNNs) especially (Dou et al., 2021; Hong et al., 2020; Khalili et al., 2019), have significantly improved quality and speed of automated segmentation. However, current models still show significant inaccuracies in certain regions that impact segmentation reliability. Here, we present a method for analyzing error patterns from CP segmentation using an attention-gated UNet model (Ronneberger et al., 2015; You et al., 2024) and compare these patterns with those derived from a novel fuzzy UNet model (Nan et al., 2022; Price et al., 2019), in order to find out clusters with largest segmentation error.

Methods: This study, approved by the Institutional Review Board at Boston Children's Hospital, retrospectively analyzed T2-weighted fetal brain MRI from 128 typically developing fetuses (gestational weeks [GW]: 29.36 ± 3.79, range: 21.86 - 37.30). 14 randomly selected cases were used for testing, while the rest were used for training segmentation models. Data was processed using our fetal brain MRI processing pipeline (You et al., 2024), which includes brain masking, non-uniformity correction, slice-to-volume registration, and alignment to the 31-week template. We then analyzed the segmentation error patterns and compared the performance of attention-gated UNet model (You et al., 2024) with a fuzzy UNet model (Nan et al., 2022; Price et al., 2019). Approach for analysis of segmentation error patterns is summarized in Figure 1. In short, we calculated average False-Positive (FP) and False-Negative (FN) maps across each image pair in the testing set. We identified ground truth edges with morphological operations and generated Distance-Error (DE) maps by calculating the Euclidean distance from each misclassified voxel to the nearest edge. Error map voxel values were then projected to the CP skeleton. We then identified the clusters with highest segmentation errors by filtering the top 50% error values on the skeleton and ranking them based on a summed error values within connected regions. Finally, we compared the within-cluster segmentation quality between attention-gated UNet and fuzzy UNet using the Dice Coefficient, Hausdorff Distance and Hybrid metrics using paired t-tests.



Results: Identified areas with the largest prevalence of segmentation errors are mainly located around Sylvian Fissure (Figure 2a&b), medial regions of the frontal lobe (figure 2a,b&c), and medial temporal lobe (Figure 2b&c). The average FP and FN errors were positively associated with GW (Linear Mixed-Effect Model, β=2221 and 2031 respectively, $p<0.001$). Comparing segmentation results between the attention-gated UNet and the fuzzy UNet, attention-gated UNet performed significantly better in terms of FP and DE only in medial regions of the frontal lobe (ΔHausdorff Distance=1.199, $p<0.05$; ΔDice=0.026, $p<0.05$, respectively), while fuzzy UNet outperformed in other metrics such as Hybrid metrics. Overall, these results suggest a comparable segmentation performance between the between the attention-gated UNet and the fuzzy UNet in the error-prone clusters.



FP-based	Dice Coefficient	Hausdorff Distance	Hybrid Metric
Region 1	(0.928, 0.926)	(6.131*, 7.330)	(0.698, 0.690)
Region 2	(0.944, 0.936)	(3.650, 3.914)	(0.733, 0.726)
Region 3	(0.921, 0.925)	(3.441, 3.773)	(0.728, 0.726)
Region 4	(0.892, 0.888)	(3.885, 3.655)	(0.693, 0.694)
Region 5	(0.912, 0.909)	(3.242, 2.639)	(0.722, 0.729)

FN-based	Dice Coefficient	Hausdorff Distance	Hybrid Metric
Region 1	(0.933, 0.930)	(6.711, 7.605)	(0.700, 0.692)
Region 2	(0.948, 0.943)	(4.909, 4.760)	(0.724, 0.721)
Region 3	(0.908, 0.901)	(5.968, 5.935)	(0.681, 0.678)
Region 4	(0.900, 0.905)	(5.315, 5.059)	(0.683, 0.690)
Region 5	(0.847, 0.828)	(3.283, 2.924)	(0.673, 0.662)

DE-based	Dice Coefficient	Hausdorff Distance	Hybrid Metric
Region 1	(0.648, 0.634)	(4.598, 4.652)	(0.515, 0.508)
Region 2	(0.501, 0.524)	(3.573, 2.989)	(0.425, 0.451)
Region 3	(0.807, 0.810)	(4.410, 4.948)	(0.644, 0.636)
Region 4	(0.740*, 0.714)	(3.175, 4.763)	(0.606, 0.573)
Region 5	(0.285, 0.311)	(3.908, 6.709)	(0.264, 0.243)

Conclusions: We characterized error patterns of automated CP segmentation. As expected, segmentation errors mainly occurred around deep sulci and low-contrast areas. This is probably due to their complex morphology, age-related brain changes, and data quality issues, such as partial volume effects in medial regions. Our segmentation error evaluation provides valuable insights for optimizing and advancing segmentation methods in the future. Building on these findings, we plan to explore advanced segmentation techniques, such as regionally weighted loss, to enhance performance in the identified clusters.

References

1. Dou, H. (2021). A Deep Attentive Convolutional Neural Network for Automatic Cortical Plate Segmentation in Fetal MRI. IEEE Transactions on Medical Imaging, 40(4), 1123–1133. https://doi.org/10.1109/TMI.2020.3046579
2. Hong, J. (2020). Fetal Cortical Plate Segmentation Using Fully Convolutional Networks With Multiple Plane Aggregation. Frontiers in Neuroscience, 14, 591683. https://doi.org/10.3389/fnins.2020.591683
3. Khalili, N. (2019). Automatic brain tissue segmentation in fetal MRI using convolutional neural networks. Artificial Intelligence in MRI, 64, 77–89. https://doi.org/10.1016/j.mri.2019.05.020
4. Nan, Y. (2022). Fuzzy Attention Neural Network to Tackle Discontinuity in Airway Segmentation (arXiv:2209.02048). arXiv. http://arxiv.org/abs/2209.02048
5. Price, S. R. (2019). Introducing Fuzzy Layers for Deep Learning. 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 1–6. https://doi.org/10.1109/FUZZ-IEEE.2019.8858790
6. Ronneberger, O. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation (arXiv:1505.04597). arXiv. https://doi.org/10.48550/arXiv.1505.04597
7. You, S. (2024). Automatic cortical surface parcellation in the fetal brain using attention-gated spherical U-net. Frontiers in Neuroscience, 18, 1410936. https://doi.org/10.3389/fnins.2024.1410936