## BiometryNet: Landmark-based Fetal Biometry Estimation from Standard Ultrasound Planes (Supplementary Material)

Netanell Avisdris  $^{\boxtimes 1,2}$ , Leo Joskowicz<sup>1</sup>, Brian Dromey<sup>4,6</sup>, Anna L. David<sup>6</sup>, Donald M. Peebles<sup>6</sup>, Danail Stoyanov<sup>4,5</sup>, Dafna Ben Bashat<sup>2,3</sup>, and Sophia Bano<sup>4,5</sup>

- <sup>1</sup> School of Computer Science and Engineering, The Hebrew U. of Jerusalem, Israel {netana03,josko}@cs.huji.ac.il
  - <sup>2</sup> Sagol Brain Institute, Tel Aviv Sourasky Medical Center, Israel
- <sup>3</sup> Sagol School of Neuroscience and Sackler Faculty of Medicine, Tel Aviv U., Israel <sup>4</sup> Wellcome/EPSRC Centre for Interventional and Surgical Sciences(WEISS), University College London, London, UK
  - Department of Computer Science, University College London, London, UK Elizabeth Garrett Anderson Institute for Women's Health, University College London, London, UK

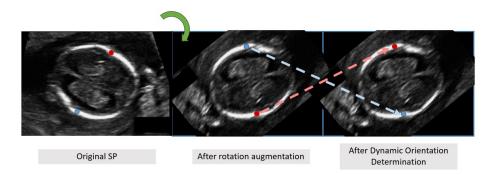
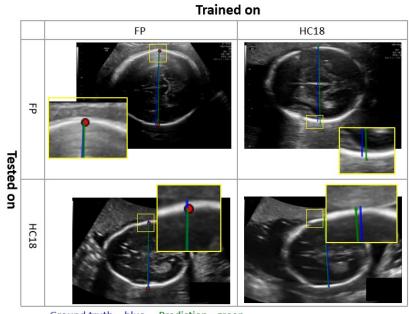


Fig. 1. Dynamic Orientation Determination (DOD): illustrative example. (a) original standard plane (SP) ultrasound (US) image of fetal trans-ventricular plane with the bi-parietal diameter (BPD) biometric measurement up (red) and down (green) ground truth landmarks; (b) SP after rotation showing an inconsistent landmark labeling in which the up and down landmarks are switched); (c) reassignment of the landmark up and down labels by DOD (blue and red intermittent lines).



Ground truth – blue Prediction - green

Fig. 2. Example showing different results of two annotation methods and the resulting BiometryNet predictions. Two models were trained, one on the FP training dataset (first column) and the second on the HC18 training dataset (second column). Both were run on the FP testing dataset (first row) and the HC18 testing dataset (second row). Landmarks are marked in the middle of the fetal skull in FP dataset and on the outer extrema of the skull in HC18 dataset. This introduces a bias that results in a relatively larger error, although still within the clinically acceptable limit when generalizing from FP $\rightarrow$ HC18 or HC18 $\rightarrow$ FP. The quantitative results are presented in Table 1 and discussed in Study 2 and 3. The images show ground truth annotations (blue) and computed predictions (green) on representative examples of BPD biometry. The relevant areas are enlarged in the yellow box.