Automated fetal brain segmentation of 2D magnetic resonance images: transfer learning and 3D topology correction

Hamza Kebiri[1,2], Priscille de Dumast[1,2], Thomas Yu[2,3], Hélène Lajous[1,2], Jean-Philippe Thiran[1,3], Reto Meuli[1], Meriam Koob[1], Meritxell Bach Cuadra[1,2,3]

1. Department of Radiology, Centre Hospitalier Universitaire Vaudois (CHUV) and University of Lausanne (UNIL)
2. Centre d'Imagerie BioMédicale (CIBM), Centre Hospitalier Universitaire Vaudois (CHUV) and University of Lausanne (UNIL)
3. Signal Processing Laboratory (LTS 5), École Polytechnique Fédérale de Lausanne (EPFL)

BACKGROUND

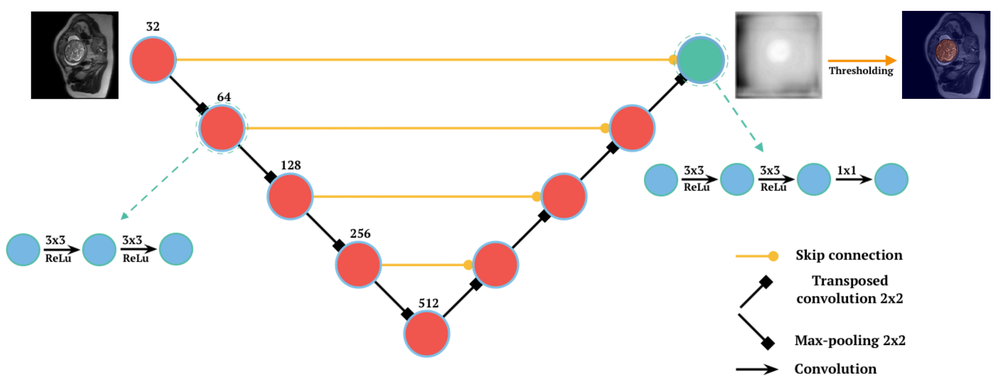
- **Fetal brain extraction in MRI**: first step for further processing (super-resolution reconstruction, tissue segmentation, etc.)
- Manual annotations are cumbersome and time consuming, and hence inappropriate to automated analysis and large-scale studies.
- **Deep learning** is an *Artificial Intelligence* branch that has proven to be very successful in image processing, including fetal brain segmentation[1-3].

Automated brain segmentation of a fetus of 26 gestational weeks

- Deep learning limitations:
  - Need of a **large amount of labeled data**
  - Highly **specialized models**
- **Transfer learning** can partially help to overcome these caveats.
- ***Aim:*** To evaluate transfer learning for segmenting the fetal brain from one dataset (Lausanne University Hospital, **CHUV**) using the pre-trained parameters of a larger dataset (Boston Children Hospital's, **BCH** [1]).

MATERIALS AND METHODS

- 2D convolutional neural network U-Net[4].



Network

- U-Net architecture[4] with ~8 million parameters
- Weighted-cross entropy loss function
- Trained with Adam optimiser for ~200 epochs

- Model evaluation in a **leave-four-out cross-validation** using an average of *precision* and *recall*

Our dataset (CHUV)

- 39 subjects
- From 20 to 36 weeks of gestation
- Orthogonal T2-weighted HASTE at **1.5T**
- 227 series totalling **4,767 slices**
- 1.125 mm in-plane isotropic
- 3 to 5 mm slice thickness

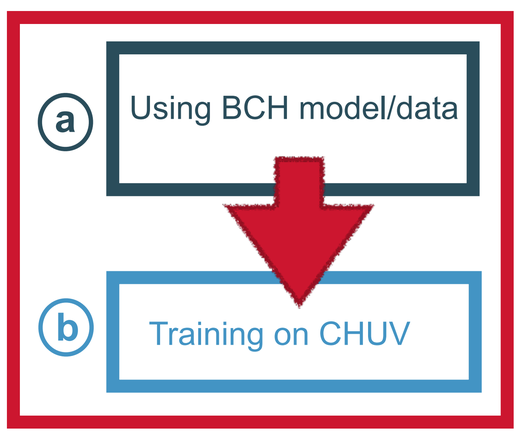
BCH dataset

- 41 subjects
- From 22 to 38 weeks of gestation
- Orthogonal T2-weighted SSFSE at **3T**
- 385 series totalling **~13,000 slices**
- 1 to 1.125 mm in-plane isotropic
- 2 to 3 mm slice thickness

- We evaluate **three scenarios**:

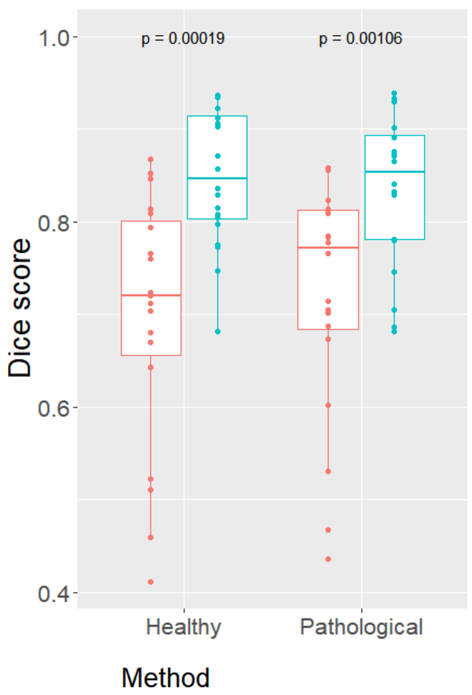
a. Using BCH pre-trained parameters[1] to test on CHUV dataset;
b. Solely training on CHUV data, with **random initialization**;
c. Fine-tuning the network with the **pre-trained** parameters of [1].

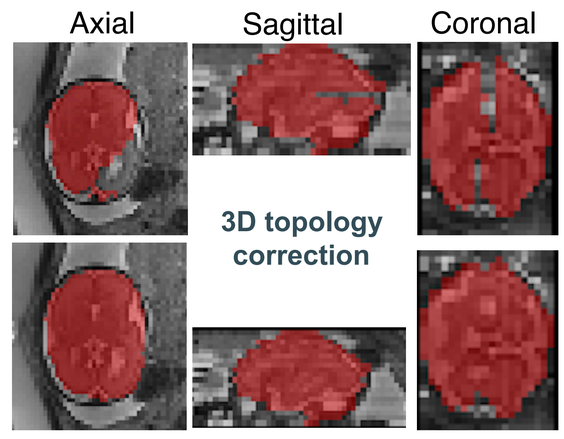
- Post-processing: **3D continuity** of the brain to refine the predictions by **morphological operations** such as closing, opening and connected components.



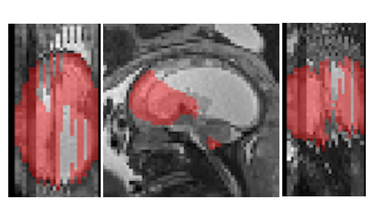
RESULTS

- Directly applying the pre-trained weights (a) from [1] to our dataset generated non plausible segmentations.
- The pre-trained network (c) significantly outperforms the randomly initialized (b) network in both healthy and pathological subjects (Wilcoxon test, $p<0.05$).
- Remaining errors: 1) at extremities of the brain, 2) slices containing the temporal lobe.
- The 3D topology correction did help qualitatively but not quantitatively.

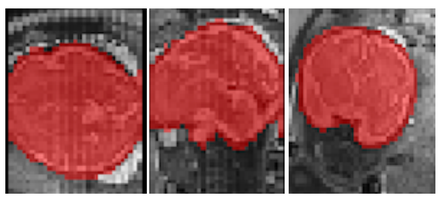




Underperformance (ventriculomegaly)



Good performance (healthy)



CONCLUSION

- Feasibility of using a different scanner/magnetic field strength through transfer learning.
- Hospitals lacking of a large amount of data can benefit from pre-trained parameters from other hospitals to boost their models.

REFERENCES: [1] Salehi et al., ISBI 2018; [2] Lou et al., MLMI 2019; [3] Rutherford et al., bioRxiv 2019; [4] Ronneberger et al., MICCAI 2015.
ACKNOWLEDGMENTS: This work is supported by the Swiss National Science Foundation (FNS projects 205321_141283 & 205321_182602) and the Hasler Foundation (17029).