Abstract title: Optimization of a deep learning strategy for estimation of cortical porosity map from MRI T1-weighted images

Author List: Matthieu Dagommer[1,2], Bastien Guerin[2,3]

[1]ESPCI, Paris France
[2]Harvard Medical School, Boston MA USA
[3]Massachusetts General Hospital/A. A. Martinos Center for Biomedical Imaging, Charlestown MA USA

Target Audience: Engineers and scientists interested in brain ultrasound for neuromodulation and high-intensity applications.

Purpose: The goal of this study is to develop an optimized deep learning strategy for estimation of cortical bone porosity maps from T1-weighted MR images for subject-specific ultrasound simulation while avoiding CT radiation exposure.

Introduction: The skull represents a barrier to ultrasound (US) propagation and penetration into the brain because of the presence of small marrow-filled pores that act as point scatterers. The resulting net effect is both absorption of the US energy and distortion of the beam, i.e. loss of focus and shape. At a minimum, US scattering should be predicted accurately in order to ensure proper dose delivery. The governing metric for this effect is cortical bone porosity, which CT scans provide however at the cost of increased exposure to ionizing radiation that is typically difficult to justify in research studies. Several studies [4], [8], [5] have shown that it is possible to obtain a good estimate of the CT scan from an MRI image using deep neural networks because, although bone is not visible in most MR sequences, the marrow is. However, skull/marrow/dura/brain MRI contrast is complex, and deep learning (DL) is a method of choice to reconstruct both the shape of the skull as well as porosity variations within that shape. We build on previous work in this area and propose a DL strategy for porosity estimation from T1w MR images that is systematically optimized for simplicity and performance. One of our original contributions is so-called "backpropagation in the mask", which we found to significantly improve learning performance.

Methods: We systematically evaluated several deep learning tactics to yield the simplest, most accurate estimation approach. *Network*: We evaluated ResNet and Unet networks embedded in a conditional Generative Adversarial Networks (cGAN) architecture. *Training*: The training data consisted in 1-mm isotropic T1-weighted MR images and CT scans acquired on 13 different subjects and aligned in the same space. Image dataset was augmented with different techniques including 90° rotations, horizontal flip operation and also by picking images from 1 to 3 orientations (sagittal, coronal and transverse). Skull masks were estimated from the MR image using Samseg [1], [2], [7] which was used as additional input information to the network. We evaluated several cropping strategies of the data: Full field-of-view (FOV), 2D patches and 3D patches of different sizes; as well as several loss functions (L1, L2, Gradient Different Loss, Perceptual Loss with Discriminator, Content Loss with external VGG-16 network and systematic use of binary cross-entropy loss as the discriminator's cGAN feedback to the generator) and a strategy whereby adjustment of the network weights during backpropagation is performed only from error pixels within the mask (BIM for "backpropagation in the mask"). *Other tests*: We evaluated advanced DL approaches including auto-context training [6] and use of the pseudo-CT [3] as additional input channel. All training was performed over 60 epochs with a Tesla P100 PCIe 16 GB GPU (48 hours), with Adam optimization (learning rate = 2.10^{-4}, $\beta_1 = 0.5$) and batch size of 5. *Validation*: This was performed in a single subject dataset.

Results: Fig. 1 shows the best strategy we identified, which consists of a cGAN architecture with a ResNet generator and PatchGAN discriminator, BIM, 32x32x32 3D-patches and a combination of L1, L2 and cGAN loss functions ($\lambda_1\mathcal{L}_{L1} + \lambda_2\mathcal{L}_{L2} + \lambda_3\mathcal{L}_{cGAN}$, with $\lambda_1 = \lambda_2 = 1$ *and* $\lambda_3 = 0.01$). 1/ Target: Using a square root of porosity target yields slightly better performance (MAE) than using porosity target (MAE reduction by 0.3%). 2/ Training set size: Fig. 2 shows that training seemed to converge after 9 training subjects, at least when using full FOV 256x256 input 2D MRI slices, indicating that the size of training set is adequate (13 subjects). 3/ Loss functions: Combining L1 and L2 training losses did not significantly reduced MAE but improved the MSE (mean square error) by 4.5% compared to when using L1 loss alone. 4/ Generators: Using a ResNet generator instead of UNet improved MAE by 25%. 5/ Patch size: 32x32x32 3D-patch models outperformed 256x256, 128x128 and 64x64 2D-patch input cropped sizes. 6/ Backpropagation: BIM improved the estimation by 11%, as assessed on the 3D-patch model and shown in Fig. 3. 7/ Miscellaneous: We found that using the pseudo-CT as a channel input, auto-context training and an additional content loss term extracted from a pre-trained VGG16 network did not improve model performance.

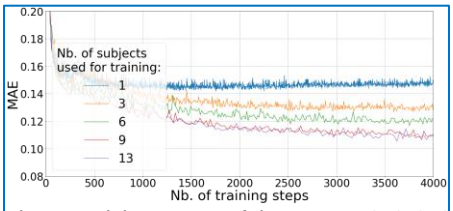

Fig. 1: Training curves of the ResNet 256x256 2D-patch model with increasing MRI/CT subject pairs in the training set. Training pairs are picked randomly with replacement, hence increasing the number of subjects in the training set increases the variability of the data that the model is exposed to during training.

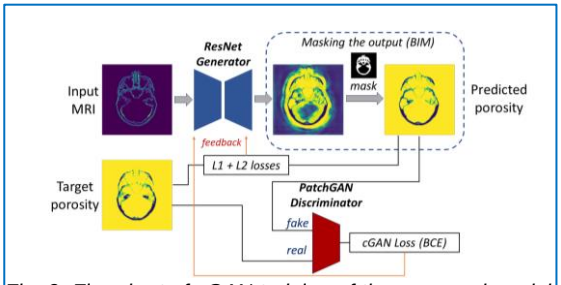

Fig. 2: Flowchart of cGAN training of the proposed model. the masked MRI input (all pixels outside the mask are 0) and the masked CT-derived porosity map output. BIM is implemented by masking the network output before backpropagation and works best in combination with masking of the input/output.

Discussion: Our proposed model integrates lessons learned from all numerical tests, and is trained on the square root of porosity instead of porosity, with L1 and L2 loss functions, uses a ResNet generator and 3D-patches and implement backpropagation in the mask (BIM), as results 1, 4, 5 and 6 suggested. Result 2 indicates that the size of the dataset is not a major limiting factor of the model performance, but of course more data is better and we are working on increasing the size of our dataset in order to build-in robustness with respect to variability in the MRI and CT contrasts due to variable acquisition parameters (such as TR, TE, CT tube energy). Fig. 3 shows that our proposed model yields excellent porosity estimations, which is because the bone marrow in the cortical bone pores is visible on the MR image. We are currently evaluating the impact of porosity estimation on ultrasound simulations.

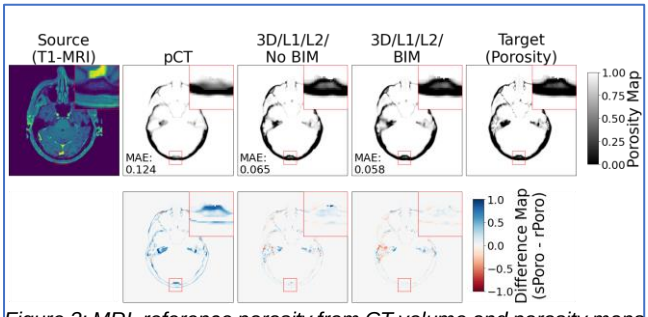

Figure 3: MRI, reference porosity from CT volume and porosity maps estimated from pseudo-CT [3] and the proposed model with and without BIM. Also shown are porosity error maps computed using the CT-derived porosity as reference. Bone marrow inside the cortical bone pores is visible on the T1w MR image (see red zoom box). pCT

Conclusion: Backpropagation in the mask is the main contribution of this study. We believe that this numerical technique can significantly improve the current porosity estimates required for acoustic simulations.

References:

[1] Stefano Cerri, Andrew Hoopes, Douglas N Greve, Mark Mühlau, and Koen Van Leemput. A longitudinal method for simultaneous whole-brain and lesion segmentation in multiple sclerosis. In *Machine Learning in Clinical Neuroimaging and Radiogenomics in Neuro-oncology*, pages 119–128. Springer.

[2] Stefano Cerri, Oula Puonti, Dominik S Meier, Jens Wuerfel, Mark Mühlau, Hartwig R Siebner, and Koen Van Leemput. A contrast-adaptive method for simultaneous whole-brain and lesion segmentation in multiple sclerosis. 225:117471.

[3] David Izquierdo-Garcia, Adam E Hansen, Stefan Förster, Didier Benoit, Sylvia Schachoff, Sebastian Fürst, Kevin T Chen, Daniel B Chonde, and Ciprian Catana. An spm8-based approach for attenuation correction combining segmentation and nonrigid template formation: application to simultaneous pet/mr brain imaging. 55(11):1825–1830.

[4] Heekyung Koh, Tae Young Park, Yong An Chung, Jong-Hwan Lee, and Hyungmin Kim. Acoustic simulation for transcranial focused ultrasound using gan-based synthetic ct. 26(1):161–171.

[5] Han Liu, Michelle K. Sigona, Thomas J. Manuel, Li Min Chen, Charles F. Caskey, and Benoit M. Dawant. Synthetic CT skull generation for transcranial MR imagingguided focused ultrasound interventions with conditional adversarial networks. In Cristian A. Linte and Jeffrey H. Siewerdsen, editors, *Medical Imaging 2022: Image-Guided Procedures, Robotic Interventions, and Modeling*. SPIE, apr 2022.

[6] Dong Nie, Roger Trullo, Jun Lian, Li Wang, Caroline Petitjean, Su Ruan, Qian Wang, and Dinggang Shen. Medical image synthesis with deep convolutional adversarial networks. *IEEE Transactions on Biomedical Engineering*, 65(12):2720–2730, dec 2018.

[7] Oula Puonti, Juan Eugenio Iglesias, and Koen Van Leemput. Fast and sequence-adaptive whole-brain segmentation using parametric bayesian modeling. 143:235–249.

[8] P. Su, S. Guo, S. Roys, F. Maier, H. Bhat, E.R. Melhem, D. Gandhi, R. Gullapalli, and J. Zhuo. Transcranial MR imagingguided focused ultrasound interventions using deep learning synthesized CT. *American Journal of Neuroradiology*, 41(10):1841–1848, sep 2020.