Realistic fMRI simulations for data augmentation in self-supervised fRMI image sequence reconstruction

Supervision:

- Philippe Ciuciu, Director of Inria-CEA MIND team (philippe.ciuciu@cea.fr, +33 1 6908 7785)
- Pierre-Antoine Comby, 3rd-year PhD candidate at Inria-CEA MIND team, NeuroSpin, (pierre-antoine.comby@ens-paris-saclay.fr)

Context. No suitable end-to-end simulator exists for modeling the functional magnetic resonance imaging (fMRI) acquisition process including first modeling brain activity in space and time on a realistic numerical template, second simulating MRI data acquisition in k-space (spatial Fourier space in a idealized scenario), third reconstructing a sequence of fMRI images and fourth performing statistical analysis to compare different acquisition protocols on a downstream validation task. Indeed, most of existing fMRI simulators proceed in the image domain by adding various sources of noise, head movements, and different types of BOLD responses, like fMRISim [1], neuRosim [2], SimTB [3], or our in-house software PyHRF[4]. In contrast, others rely on MR physics by generating the raw MRI signal from the input parameters in a MR pulse sequence such as POSSUM (Physics-Oriented Simulated Scanner for Understanding MRI) [5] or ODIN [6]. However, the latter are not really specific to fMRI.

To address these multiple issues and offer a complete approach, we have recently introduced a new, modular, open-source, python-based, fMRI simulation software called `snake-fmri`. Unlike existing tools, the goal is here to simulate the whole acquisition process of fMRI signals, from the evoked brain responses to the multi-coil k-space acquisition and reconstruction, with the possibility of extending the forward model to various noise and artifact sources while remaining memory-efficient. By using this *in silico* setup, we are thus able to provide realistic ground truth for fMRI reconstruction methods and explore the influence of critical parameters, such as acceleration factor and signal-to-noise ratio (SNR), on the downstream processes of fMRI image sequence reconstruction and statistical analysis of evoked brain activity.

Motivation and Objective. With the recent advances of accelerated fMRI acquisition, the time and complexity saved at the acquisition has been transferred to the image reconstruction task. As of now, even though modern variational low-rank + sparse reconstruction techniques have been developed in our software pysap-fmri, the time budget required to fully reconstruct a typical sequence of 100 volumes is too much as approximately between 3 and 10 hours on a multi-GPU architecture. Therefore, the recent trend is to leverage deep learning to address this issue, notably by unrolling optimization algorithms, as we did for structural MRI [7], and then by replacing usual denoisers with convolutional neural networks or more advanced structures (e.g. transformers). The difficulty for fMRI lies in the lack of a ground truth to apply supervised learning. Therefore in the fMRI community, the recent tendency is to move towards self-supervised learning [8], but with the well known pitfalls, such as the limited acceleration factor and relatively high SNR constraints, thereby limiting the exploration of high spatio-temporal fMRI data sets.

Objective: In this internship, the goal is therefore to address this issue by implementing data augmentation using the simulator `snake-fmri` to generate multiple transformations of an actual fMRI data set, hence robustifying usual self-supervised approaches.

Methods. The proposed internship follows a 3-stage methodology:

1. **Unrolling low-rank+sparse reconstruction algorithms.** The first step will consist in unrolling our ADMM implementation of the fMRI image reconstruction algorithm that promotes low-rank (in space) + sparse (in the

temporal Fourier domain or any appropriate transformed domain) prior information. To this end, we will leverage the open source implementation of the NCPDNet architecture and the seminal paper [9], and adapt it to fMRI data. The traditional Singular Value Decomposition (SVD) used for promoting low-rank representation of the fMRI image sequence will be transformed into a learned SVD, while the usual proximity operator that encodes sparsity will be replaced by a learned version too. These operators will be learned typically through a CNN structure involving convolution, residual and max-pooling blocks over at least 10 iterations. Data consistency blocks will be used to make sure that the denoised fMRI images are compatible with the original k-space data once put back in the k-space domain. The implementation will be performed in `PyTorch` and access granted to Jean Zay supercomputer (GPU A100 Nvidia boards) will make the testing fast enough.

2. **Preliminary testing and validation.** The candidate will generate extensive synthetic fMRI data using the `snake-fmri` simulator under various acquisition scenarios (in terms of SNR, spatial resolution, acceleration factor, and Cartesian vs non-Cartesian readout) to test and validate the proposed deep learning architecture on synthetic fMRI image sequence reconstruction.
3. **Data augmentation in self-supervised learning.** The classical way to perform self-supervised reconstruction in fMRI is to split the k-space data (sequence of volumes in the Fourier domain) in two sets using a hold out masking strategy [8], learn the network on the first half of the data and validate/test on the second half. Various data transformations (e.g., rotation, dephasing) will be applied to real fMRI data sets thanks to the possibilities offered by `snake-fmri` and the impact of this strategy will be first evaluated on image quality, and then on a downstream statistical validation task to detect evoked brain brain activity with the best statistical sensitivity/specificity compromise.

Environment. The internship will take place at NeuroSpin, in the MIND team for maximally 6 months (April-September 2024). This is a large team focused on mathematical methods for statistical modeling of brain function using neuroimaging data (fMRI, MEG, EEG) as well as advanced computational imaging methods. Particular topics of interest include machine learning techniques, numerical and parallel optimization, advanced statistical approaches in functional neuroimaging, applications to cognitive and clinical neuroscience, and scientific software research.

Skills. We seek candidates who are strongly motivated by challenging research topics in machine/deep learning and neuroimaging. Applicants should have a strong mathematical background with knowledge in numerical optimization, machine and deep learning. With regards to software engineering, proficiency in `Python` is expected and strong experience in a deep-learning library such as `PyTorch` is necessary.

References

[1] Ellis, C. T., Baldassano, C., Schapiro, A. C., Cai, M. B. & Cohen, J. D. Facilitating open-science with realistic fMRI simulation: validation and application. *PeerJ* **8**, e8564 (2020).

[2] Welvaert, M., Durnez, J., Moerkerke, B., Verdoolaege, G. & Rosseel, Y. neuRosim: An R package for generating fMRI data. *Journal of Statistical Software* **44**, 1–18 (2011).

[3] Erhardt, E. B., Allen, E. A., Wei, Y., Eichele, T. & Calhoun, V. D. SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability. *Neuroimage* **59**, 4160–4167 (2012).

[4] Vincent, T. *et al.* Flexible multivariate hemodynamics fMRI data analyses and simulations with PyHRF. *Frontiers in Neuroscience* **8**, 67 (2014).

[5] Drobnjak, I., Pell, G. S. & Jenkinson, M. Simulating the effects of time-varying magnetic fields with a realistic simulated scanner. *Magnetic resonance imaging* **28**, 1014–1021 (2010).

[6] Jochimsen, T. H., Schäfer, A., Bammer, R. & Moseley, M. E. Efficient simulation of magnetic resonance imaging with Bloch–Torrey equations using intra-voxel magnetization gradients. *Journal of Magnetic Resonance* **180**, 29–38 (2006).

[7] Ramzi, Z., Chaithya, G., Starck, J.-L. & Ciuciu, P. NC-PDNet: A density-compensated unrolled network for 2D and 3D non-Cartesian MRI reconstruction. *IEEE Transactions on Medical Imaging* **41**, 1625–1638 (2022).

[8] Singh, D. *et al.* Emerging Trends in Fast MRI Using Deep-Learning Reconstruction on Undersampled k-Space Data: A Systematic Review. *Bioengineering* **10**, 1012 (2023).

[9] Huang, W. *et al.* Deep low-rank plus sparse network for dynamic MR imaging. *Medical Image Analysis* **73**, 102190 (2021).