3D stitching of ultra-terabyte brain images and feature extraction with deep learning

Giacomo Mazzamuto[1], Francesco Orsini[2], Matteo Roffilli[3], Paolo Frasconi[2], Francesco Saverio Pavone[1,4,5], Ludovico Silvestri[1,5]
[1]European Laboratory for Non-linear Spectroscopy (LENS), Sesto Fiorentino, Italy
[2]Department of Information Engineering (DINFO), University of Florence, Italy
[3]Bioretics Srl, Cesena, Italy
[4]Department of Physics and Astronomy, University of Florence, Italy
[5]National Institute of Optics (INO), National Research Council (CNR), Sesto F.no, Italy

Email: mazzamuto@lens.unifi.it

Investigating the brain architecture at a system level is a demanding task, which requires the ability to quantitatively map the brain structure with cellular and subcellular resolution. Besides posing significant challenges to current optical microscopy methods, a new generation of software tools needs to be developed to make sense of the huge amount of raw images generated, which can easily exceed several TeraBytes for a single sample (e.g. whole mouse brain tomographies acquired with a Light-Sheet Fluorescence Microscope). Here we present an image processing pipeline that enables us to extract a semantic representation of the sample starting from a collection of 3D tiles of raw gray levels as produced by the LSFM.

As a first step, the hundreds of adjacent and overlapping tiles produced by the microscope need to be aligned and fused together. To this aim, we developed ZetaStitcher [1], a Python package for volumetric stitching that computes global optimal alignment of imaging datasets as big as 8 TB in less than an hour. The fused volume is then accessed virtually, without the need to create a physical copy of the dataset, by means of a dedicated API.

In order to extract meaningful information, the reconstructed virtual volume is queried using the API and fed in chunks to other pieces of software for image analysis. We demonstrate two complementary approaches based on deep convolutional neural networks. In one case, a 3D conv-net is used to perform "semantic deconvolution" of the image [2], enabling accurate localization of neuronal bodies with standard clustering algorithms (e.g. mean shift). The scalability of this approach is demonstrated by mapping the spatial distribution of different neuronal populations with single-cell resolution in a whole mouse brain.

To go beyond simple localization, in samples of human brain cortex imaged with a Two-Photon Fluorescence Microscope, we exploited a 2D conv-net estimating the probability for each pixel to be the center of a patch containing the visual pattern of a neuron [3]. The probabilistic maps produced by the CNN are then processed with a contour finding algorithm, obtaining reliable segmentation of cell morphology.

[1] G. Mazzamuto. https://github.com/lens-biophotonics/ZetaStitcher
[2] P. Frasconi, L. Silvestri, P. Soda, R. Cortini, F. S. Pavone and G. Iannello, "Large-scale automated identification of mouse brain cells in confocal light sheet microscopy images", Bioinformatics 30(17), i587-i593 (2014)
[3] G. Mazzamuto, I. Costantini, M. Neri, M. Roffilli, L. Silvestri, F. S. Pavone, "Automatic segmentation of Neurons in 3D Samples of Human Brain Cortex", Applications of Evolutionary Computation. EvoApplications 2018. Lecture Notes in Computer Science, vol 10784. Springer 2018