Computing for Optimized Biological Microscopy Data Processing and Analysis at The Rosalind Franklin Institute

Luís M. A. Perdigão[1*], Neville B.-y. Yee[1], Elaine M. L. Ho[1], Avery H. Pennington[2], Oliver N.F. King[2], Michele C. Darrow[1], Mark Basham[1]

[1.] The Rosalind Franklin Institute, Harwell Campus, Didcot, United Kingdom.
[2.] Diamond Light Source, Harwell Campus, Didcot, United Kingdom.
* Corresponding author: luis.perdigao@rfi.ac.uk

Recent developments in experimental microscopy techniques have led to improvements in the way we visualize various biological phenomena. Presently, state-of-art microscopy involves cryogenic sample preparation, 3D correlative microscopy and milling, followed by tilt series acquisition of biological volumes leading to datasets with nanometer scale information. While this workflow is technically possible, it is still challenging to collect, process, and analyze these large datasets, especially when the workflow includes correlative imaging and segmentation steps. In the Artificial Intelligence and Informatics group (AI&I) at The Rosalind Franklin Institute we are automating these workflow steps to solve computationally difficult and time-intensives problems by developing open-source software tools. Here, we present some notable examples.

RedLionfish [1,2] is a package for fast GPU/CPU accelerated Richardson-Lucy deconvolution of 3D optical images. This is useful for removing optical artifacts in 3D microscopy data, leading to clearer and sharper images. This software is available as a plugin for napari [3] (a 3D data visualization application) and it has been included in 3DCT [4] as a data processing tool for correlative microscopy, including speed improvements such that it can be used in real time alongside of focus ion beam (FIB) lamella preparation (Figure 1A).

Ot2Rec [5,6] is a package for automating the tomographic reconstruction workflow. This software is a wrapper for the commonly used processing packages MotionCor2 (motion correction) [7], CTFFind4 (CTF estimation) [8], IMOD [9] and Savu [10] (alignment and reconstruction). Although processing pipeline solutions already existed (EMAN2, tomoBEAR), this solution offers advantages through a more general, unified command-line syntax, flexibility for future expansion and a more portable codebase. In addition to the above processing packages, Ot2Rec also includes a tool for simulating CTF image stacks and generating 3D point spread functions (PSF) for deconvolution tasks (Figure 1B).

SuRVoS2 [11-13] is a collection of tools to help accelerate annotation and segmentation of large volumetric bio-imaging workflows. It enables either shallow or deep machine learning approaches, using a suite of image processing filters, supervoxels (boundary adherent groupings of similar, adjacent voxels), and annotation hierarchies. SuRVoS2 also provides a set of tools to enable visualization and interaction with large numbers of distributed annotations (e.g. performed by multiple members of a group or citizen scientists). This application has been implemented both as a multi-platform napari plugin and as an API for generic programming usage.

Unet+ [14,15] is a new approach to machine learning training and prediction for segmentation of large volume biological samples that uses the well-known UNet [16] neural network architecture for 2D

biological images, but expands to allow predictions of 3D volumes, using a multi-slicing, multi-axis and multi-rotation technique. Computational post-processing methods are currently being developed that combine the generated multi-predictions for optimized confidence metrics.

As tomography experiments become increasingly more technically complex, future software development needs to follow up by aiming for improving data processing pipelining tasks and maintaining accessibility by being open-source and integration with commonly used visualization packages, as the ones presented here.
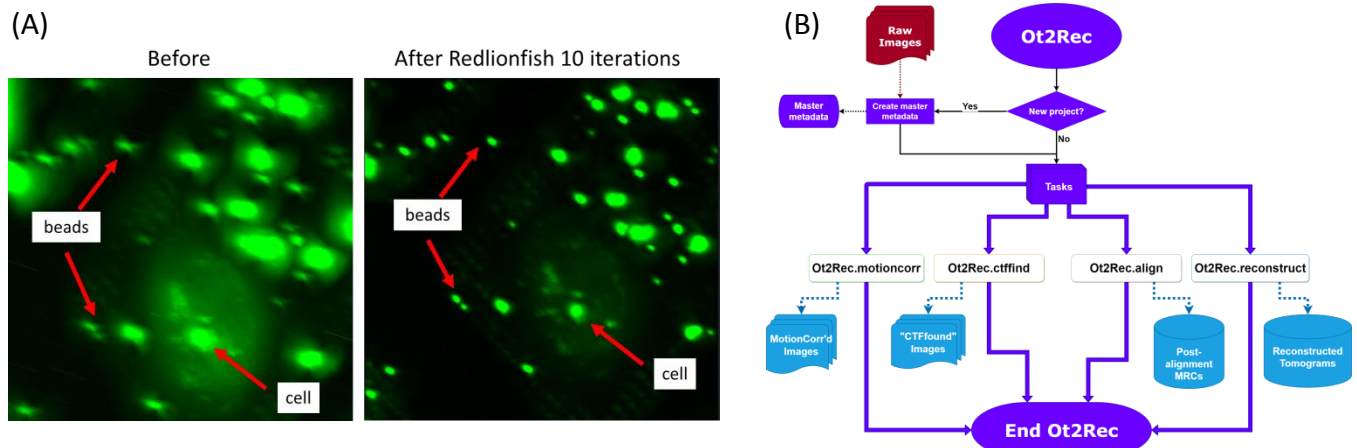
Figure 1. (left) Redlionfish example in deconvolution of 3D fluorescence data of a beaded cell sample. The deconvolution allows for more precise targeting during FIB-milling. (B) Usage workflow of Ot2Rec wrapper for automated processing of tomographic data.
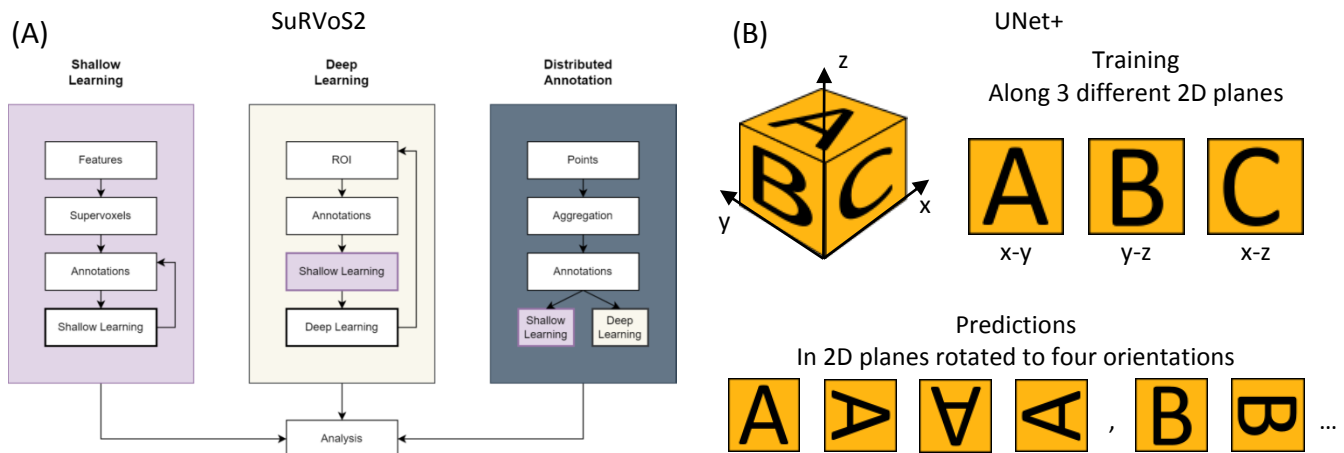

Figure 2. (A) Summary of SuRVoS2 capabilities for shallow and deep learning implementations of semi-automatic segmentation strategies, and distributed annotation tools for working with large numbers of annotations of biological data volumes. (B) Diagram of the Unet+ methodology for training and predictions augmentations, illustrating 2D slicing over different planes and different rotations.

[1] LMA Perdigao et al., manuscript in preparation.

[2] Red lionfish (RL) deconvolution. Rosalind Franklin Institute,
https://github.com/rosalindfranklininstitute/RedLionfish (accessed: Feb. 16, 2022).

[3] napari contributors (2019). napari: a multi-dimensional image viewer for python.
doi:10.5281/zenodo.3555620.

[4] J Arnold et al., Biophysical Journal, **110**, 4 (2016) p. 860, doi: 10.1016/j.bpj.2015.10.053.

[5] Neville B.-y. Yee, E Ho and M Basham, manuscript in preparation.

[6] Ot2Rec. Rosalind Franklin Institute, https://github.com/rosalindfranklininstitute/Ot2Rec (accessed:
Feb. 16, 2022).

[7] S Zheng et al., Nat Methods **14**, 4, 4 (2017), p. 331. doi: 10.1038/nmeth.4193.

[8] A Rohou, N Grigorieff, J Struct Biol. **192(2)** (2015) p. 216-21. doi: 10.1016/j.jsb.2015.08.008.

[9] JR Kremer, DN Mastronarde and JR McIntosh, J. Struct. Biol. **116** (1996) p. 71. doi:
10.1006/jsbi.1996.0013.

[10] N Wadeson and M Basham, arXiv preprint (2016) arXiv:1610.08015.

[11] A. Pennington et al., submitted.

[12] M. C. Darrow et al., JoVE (Journal of Visualized Experiments), **126** (2017) p. e56162. doi:
10.3791/56162.

[13] SuRVoS2. Diamond Light Source, https://github.com/DiamondLightSource/SuRVoS2 (accessed:
Feb. 16, 2022).

[14] WM Tun et al., Journal of The Royal Society Interface, **18,** 179 (2022) p. 20210140. doi:
10.1098/rsif.2021.0140.

[15] ONF King et al., manuscript in preparation.

[16] O Ronneberger, P Fischer and T Brox, in Medical Image Computing and Computer-Assisted
Intervention – MICCAI 2015, N Navab, J Hornegger, WM Wells, AF Frangi, Eds. (Springer
International Publishing, Cham, **vol. 9351 of Lecture Notes in Computer Science** (2015) p. 234. doi:
10.1007/978-3-319-24574-4_28.