

CREATING USER FRIENDLY SOLUTIONS FOR BIG DATA PROCESSING

Hoang Anh Nguyen¹, Zane van Iperen¹, Oliver Cairncross¹, Jake Carroll¹, David Abramson¹, James Springfield², Nicholas D Condon²

¹Research Computing Centre (RCC)

²Institute for Molecular Bioscience

The University of Queensland, Queensland, Australia

Email: n.condon@uq.edu.au

Twitter: @DrNickCondon

KEY WORDS: Big Data, Image Processing, Deconvolution, File Conversion, HPC, Cluster

ABSTRACT

Rates of acquisition and data generation are ever increasing, and microscopists can now easily produce terabytes of data in single imaging sessions. Scientists and imaging facilities can quickly become overwhelmed with dealing with this glut of data, with challenges including, data migration, storage & archiving, processing, and visualising these large datasets. Recently we have developed and deployed tools and workflows to better handle big microscopy data for our researchers, including a parallel file system with local caches, Virtual desktops (Windows & Linux) for image processing, and a website based HPC job submission platform called the *Image Processing Portal* [1].

The screenshot displays the 'IMB Image Processing Portal' interface. At the top, there is a navigation bar with the university logo and menu items: Portal, FileManager, Converter, Preprocessing, Deconvolution, JobList, Support, and Logout. Below this, the 'Light Sheet' section is active, showing a 'MAIN' tab with various configuration options. These include 'Lateral spacing (nm/pixel)' set to 100, 'Axial spacing (nm/slice)' set to 312, and 'PSF Model' set to 1.4. There are also checkboxes for 'Generate PSF' and 'Read spacing from metadata'. The interface includes a file selection area on the left and a 'SUBMIT' button at the bottom. The footer contains logos for RCC, IMB, and TERN.

The Image Processing portal contains file management tools, file conversion plugins, image pre-processing tools and deconvolution using Microvolution [2], all of which run on organisational HPC infrastructure.

Additionally, interactive tools using ImageJ/FIJI have been prepared for bulk processing of data in user-friendly ways, including bulk Z-projection scripts, for fluorescence and or brightfield images, as well as dimensional de-stacking (to chop up individual files by time, Z-stack, channel etc) [3].

[1] <https://imbmicroscopy.rcc.uq.edu.au>

[2] <https://microvolution.com/>

[3] <https://github.com/nickcondon>