Galaxy as an Integration and Workflow Platform for Bio-medical Image Analysis and Image Processing Toolkit

Piotr Szul, Dadong Wang, Yulia Arzhaeva, Shiping Chen, Alex Khassapov, Neil Burdett, Timur Gureyev, John Taylor, Tomasz Bednarz

Cloud  50,000 concurrent tasks  
Data  Scalable @ 10-20 PB  
Data  @ 1-2 PB  
HPC  Capability @ > 1 Pf/s  
HPC  Specialised @ 100 Tf/s  
Network Layer-1 @ Nx100 Gb/s  
Network Layer-3 @ 10 Gb/s  

NeCTAR Research Cloud...  

The Australian eResearch (Backbone) Infrastructure @ 2013  

40 Software Tools and Virtual Laboratory Projects  
250 Data improvement projects @ all institutions  

NeCTAR – slides by Dr Nigel Ward of NeCTAR
NeCTAR is funding four programs

**Research software**
- Virtual laboratories
- eResearch Tools

**Computational platforms**
- Research Cloud
- National Server Program
Project Vision

We have 2-3 / 3D data coming from different image modalities

**Need:** image analysis, image processing, image reconstruction tools

**Need:** fast, coherent cloud-based image analysis & processing tools integrating existing CSIRO software packages:

- HCA-Vision, X-TRACT, MILXView,
- Workspace.
HCA-Vision

Developed by CSIRO Quantitative Imaging group for automating process of quantifying cells features in microscopy images. It can reproducibly analyse complex cell morphologies. Recently, extended to 3D enables the analysis of neuron structures in vitro (in cells cultured in a 3D gel matrix) and in vivo (e.g. in exposed rat brains or viable rat brain tissue sections).

It has great value in particular for the pharmaceutical and neuroscience research community.

**Figure:** Neurite Analysis (a) input image; (b) resulted image; (c) diagram of the algorithm.
X-TRACT

A software for advanced X-ray image analysis and Computed Tomography currently in use on the MASSIVE cluster at the Australian Synchrotron, ANU and at the Shanghai Synchrotron in China.

X-TRACT implements a large number of conventional and advanced algorithms for 2D and 3D X-ray image reconstruction and simulation.

Figure: (a) Insect, reconstruction and rendering by Sherry Mayo (CSIRO); (b) Acacia plant, sample (~1 mm across) provided by Mel Linton (CSIRO), collected, reconstructed and rendered by Sherry Mayo; (c) Sample input Sinogram.
MILXView

A 3D medical imaging analysis and visualisation platform increasingly popular with researchers and medical specialists working with MRI, PET and other types of medical images.

Figure: (a) Brain tumor - PET scan and MRI overlaid; (b) CT scan of a prostate of a patient overlaid with radiation dose; (c) Generated 3D view of a brain allowing study of atrophy pattern characteristics of diseases such as Alzheimer's disease.
Galaxy as an Integration and Workflow Platform for...
Piotr Szul, Tomasz Bednarz, Dadong Wang, Yulia Arzhaeva, Shiping Chen, Neil Burdett, Alex Khassapov & Luke Domanski
Glue = Galaxy
Extract 2D Slices

Extracts 2D slices in three directions from a 3D image
Input image $\leftarrow$ 138_T1_MRI.nii.gz
It takes few seconds and a HTML is created to show the user the 2D slices and also lets the user download the images
Visualisation

WebGL based, Slice:Drop
Cellular imaging in Galaxy
Building Astrocytes analysis workflow

Step 1. Find nuclei
Building Actrocytes analysis workflow

Output: nuclei binary mask
Building Astrocytes analysis workflow

Step 2. Find lines

![Image of Galaxy tool showing find lines function]

- Input image: Control_EAAT1_n15a_rgb.tif
- Line length: 9
- Line length:
- Ignore objects smaller than: 5
- Small objects max length: 2.06
- Line sensitivity: Contract
- Smoothing size: 5
- Smoothing element size:
- Link Distance:
- Gap size between links: 10.0
- Link quality: 50
- Percentage of intensity drop in gaps
- Ignore lines with intensity lower than: 10.0
- Line intensity threshold
- Select image channel: Green channel
- Suitable image channel

This component finds lines in an image.
Building Astrocytes analysis workflow

Output: lines binary mask
Building Astrocytes analysis workflow

Last step. Compute statistics on lines per each cell

```
This compute statistics, such as line length, angle, density etc, for lines found in the image.
```
Building Actrocytes analysis workflow

Output: lines statistics
Complete Astrocytes analysis workflow

Astrocytes analysis workflow can be reused with other image data
NeCTAR Imaging Toolkit Production Deploy

The following job has been successfully added to the queue:
9: CT Reconstruction on sime_0001.qrd at Fri May 10 05:03:54 2013

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'Running' to 'Finished' if completed successfully or 'Error' if problems were encountered.

CloudMan from Galaxy
Welcome to CloudMan. This application allows you to manage your instance, cloud cluster and the services provided within. Your previous data store has been reconstructed. Once the cluster has initialized, use the control below to manage services provided by the application.
Where are we ...

So far:

• Migrated all the packages to Linux platform
• Defined domain specific data types and developed most of the tools
• Small scale cloud deployments for pilot users
• Prepared training materials for the users
• Gathered initial feedback from the user communities

Next steps:

• Full scale deployment on the Research Cloud
• Bringing more users on and ongoing improvement of the platform
• Create a ToolShed and refactor the toolkit
• Develop and share workflows
Cloud-Based Image Analysis and Processing Toolbox

HCA-VISION

MILXVIEW

X-TRACT

SAMPLE GALAXY WORKFLOW

Research community
Workflows
Clouds
Visualisation
Collaboration
Bio-Imaging
Big Data Platform

PROJECT TEAM


http://cloudimaging.blogspot.com.au

Galaxy as an Integration and Workflow Platform for ... | Piotr Szul, Tomasz Bednarz, Dadong Wang, Yulia Arzhaeva, Shiping Chen, Neil Burdett, Alex Khassapov & Luke Domanski
Thank you

CMIS / CSS TCP
Piotr Szul, Senior Software Engineer

t +61 2 xxxx xxx
E piotr.szul@csiro.au
w www.csiro.au/cmis
# HCA-Vision High Level Functionality

<table>
<thead>
<tr>
<th>ID</th>
<th>FUNCTION</th>
<th>SHORT DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>H.01</td>
<td>Detect nuclei</td>
<td>Detect nuclei in a 2D microscope image</td>
</tr>
<tr>
<td>H.02</td>
<td>Detect nuclei from cytoplasm holes</td>
<td>Detect nuclei from absence of stain of cytoplasm</td>
</tr>
<tr>
<td>H.03</td>
<td>Detect cells with nuclei</td>
<td>Detect cells using nucleus image as a mask</td>
</tr>
<tr>
<td>H.04</td>
<td>Detect cells without nuclei</td>
<td>Detect cells without using nucleus image as a mask</td>
</tr>
<tr>
<td>H.05</td>
<td>Detect neurons with nuclei</td>
<td>Detect neurons from a neurite outgrowth image using nucleus image as a mask</td>
</tr>
<tr>
<td>H.06</td>
<td>Detect neurons without nuclei</td>
<td>Detect neurons from a neurite outgrowth image without using nucleus image as a mask</td>
</tr>
<tr>
<td>H.07</td>
<td>De-clump touching objects</td>
<td>Separate any touching objects in an image, such as touching nuclei or cells</td>
</tr>
<tr>
<td>H.08</td>
<td>Label objects</td>
<td>Label objects such nuclei or cells in a binary image</td>
</tr>
<tr>
<td>H.09</td>
<td>Get object stats</td>
<td>Retrieve statistical features of individual objects in a segmented image, including area, perimeter, origin, width and height of the bounding box, coordinate of the centroid, major and minor axis of best fit ellipse, approximate of area of convex hull etc.</td>
</tr>
<tr>
<td>H.10</td>
<td>Detect cell from nucleus donuts</td>
<td>Get an anisotropic doughnut which is a region around a cell nucleus that is not uniformly thick. The extension of the doughnut is larger along the major axis of the nucleus than perpendicular to it.</td>
</tr>
<tr>
<td>H.11</td>
<td>Detect dots</td>
<td>Detect dots in a 2D image or a cell</td>
</tr>
<tr>
<td>H.12</td>
<td>Detect lines</td>
<td>Detect line structures in a 2D image or a cell</td>
</tr>
<tr>
<td>H.13</td>
<td>Get dot stats</td>
<td>Retrieve statistical features of the detected dots, including area, perimeter etc.</td>
</tr>
<tr>
<td>H.14</td>
<td>Get line stats</td>
<td>Retrieve the statistical features of detected line structures</td>
</tr>
<tr>
<td>H.15</td>
<td>Cell Scoring</td>
<td>Count negative and positive cells, measure integrated and average intensity of negative and positive cells.</td>
</tr>
</tbody>
</table>

**IDs H.xx – Application area: cell features of microscopy images**
## X-TRACT High Level Functionality

<table>
<thead>
<tr>
<th>ID</th>
<th>FUNCTION</th>
<th>SHORT DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>XP.01</td>
<td>Sinogram creation</td>
<td>X-ray projection data must first be converted into sinograms before CT reconstruction can be carried out. Each sinogram contains data from a single row of detector pixels for each illuminating angles. This data is sufficient for the reconstruction of a single axial slice (at least, in parallel-beam geometry).</td>
</tr>
<tr>
<td>XP.02</td>
<td>Ring artefact removal</td>
<td>Ring artefacts are caused by imperfect detector pixel elements as well as by defects or impurities in the scintillator crystals. Ring artefacts can be reduced by applying various image processing techniques on sinograms or reconstructed images.</td>
</tr>
<tr>
<td>XP.03</td>
<td>Dark current subtraction</td>
<td>Dark current subtraction compensates for the readout noise, ADC offset, and dark current in the detector. The dark current images are collected before and/or after CT measurements with no radiation applied and with the same integration time as the one used during the measurements. The dark current image is subtracted from each CT projection.</td>
</tr>
<tr>
<td>XP.04</td>
<td>Flat field correction</td>
<td>Flat-field images are obtained under the same conditions as the actual CT projections, but without the sample in the beam. They allow one to correct the CT projections for the unevenness of the X-ray illumination.</td>
</tr>
<tr>
<td>XP.05</td>
<td>Positional drift correction</td>
<td>The function is used for correction of transverse drift between related experimental images. Image drift is assessed by cross-correlating pairs of images.</td>
</tr>
<tr>
<td>XP.06</td>
<td>Data normalisation</td>
<td>Data normalisation including normalisation to a user-defined region</td>
</tr>
<tr>
<td>XP.07</td>
<td>TIE-based phase extraction</td>
<td>The TIE algorithm allows the recovery of the optical phase of an electromagnetic wave (e.g. an X-ray beam) from a single near-field in-line image by solving the Transport of Intensity equation under the assumption that the phase shift and absorption distributions are proportional to each other. This method is usually applied in propagation-based in-line CT imaging (PCI-CT).</td>
</tr>
<tr>
<td>XCT.01</td>
<td>FBP CT reconstruction</td>
<td>Filtered back-projection (FBP) parallel-beam CT reconstruction</td>
</tr>
<tr>
<td>XCT.02</td>
<td>FDK CT reconstruction</td>
<td>Feldkamp-Davis-Kress (FDK) cone-beam CT reconstruction</td>
</tr>
<tr>
<td>XCT.03</td>
<td>Centre of rotation</td>
<td>Automated calculation of the centre of sample rotation in a CT scan from experimental X-ray projections, sinograms or reconstructed axial slices.</td>
</tr>
<tr>
<td>XCT.04</td>
<td>CT Reconstruction Filters</td>
<td>The choice of available CT reconstruction filters will include at least the Liner-Ramp, Shepp-Logan, Cosine, Hamming and Hann filters.</td>
</tr>
<tr>
<td>XCT.05</td>
<td>ROI reconstruction</td>
<td>This option enables the user to select a subset of axial slices to be reconstructed and/or limit the reconstruction area to a user-defined rectangular subarea of the axial slice. The option reduces the reconstruction time and the size of the output data.</td>
</tr>
</tbody>
</table>

IDs **XP.xx** – Application area: **data processing functions**
IDs **XCP.xx** – Application area: **CT reconstruction functions**
## MILXView High Level Functionality

<table>
<thead>
<tr>
<th>ID</th>
<th>FUNCTION</th>
<th>SHORT DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC.01</td>
<td>Atlas registration</td>
<td>Align an atlas image to a target image</td>
</tr>
<tr>
<td>MC.02</td>
<td>Segmentation</td>
<td>Segment the MRI into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF)</td>
</tr>
<tr>
<td>MC.03</td>
<td>Bias Field Correction</td>
<td>Estimate and remove the noise on the image</td>
</tr>
<tr>
<td>MC.04</td>
<td>Partial Volume estimation</td>
<td>Quantify the amount of partial voluming inside each voxel</td>
</tr>
<tr>
<td>MC.05</td>
<td>Topology Correction</td>
<td>Create the topology of the brain to ensure that it is genus zero</td>
</tr>
<tr>
<td>MC.06</td>
<td>Thickness Estimation</td>
<td>Compute the thickness of the cortex for each Grey matter voxel</td>
</tr>
<tr>
<td>MS.01</td>
<td>Cortical surface extraction</td>
<td>Extract a 3D mesh from the brain segmentation</td>
</tr>
<tr>
<td>MS.02</td>
<td>Topological correction</td>
<td>Remove holes and handles from the mesh</td>
</tr>
<tr>
<td>MS.03</td>
<td>Biomarker mapping on cortical surface</td>
<td>Mapping of various values on the mesh i.e. thickness, PET values, MR intensity etc ...</td>
</tr>
<tr>
<td>MS.04</td>
<td>Surface registration</td>
<td>Align the meshes of any given subject to a template to obtain a correspondence across subjects</td>
</tr>
<tr>
<td>MS.05</td>
<td>Transfer of biomarkers on template surface</td>
<td>Map all the values from all subjects to a common space where they can be compared</td>
</tr>
<tr>
<td>MP.01</td>
<td>PVC Registration</td>
<td>Registration of the PET image to its corresponding MRI</td>
</tr>
<tr>
<td>MP.02</td>
<td>Segmentation</td>
<td>Segmentation of the MRI into GM, WM, and CSF</td>
</tr>
<tr>
<td>MP.03</td>
<td>Partial Volume correction (PVC)</td>
<td>Correction for spill in and spill over of the PET image using the MRI segmentation</td>
</tr>
<tr>
<td>MR.01</td>
<td>SUVR Registration</td>
<td>Registration of the PET image to its corresponding MRI</td>
</tr>
<tr>
<td>MR.02</td>
<td>Segmentation</td>
<td>Segmentation of the MRI into GM, WM and CSF</td>
</tr>
<tr>
<td>MR.03</td>
<td>Atlas Registration</td>
<td>Registration of an atlas to the MRI to define a reference region on the MRI</td>
</tr>
<tr>
<td>MR.04</td>
<td>Image Normalisation</td>
<td>Normalising the PET intensity with the intensity of the reference region</td>
</tr>
</tbody>
</table>

**IDs**
- **MC.xx** – Application area: neuro-imaging analysis, cortical thickness estimation (CTE)
- **MS.xx** – Application area: neuro-imaging analysis, CTE surface
- **MP.xx** – Application area: neuro PET analysis, PET PVC
- **MR.xx** – Application area: neuro PET analysis, PET SUVR