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

COMPUTATIONAL SIMULATION SCIENCES TCP www.csiro.au



https://www.nectar.org.au/cloud-based-image-analysis-and-processing-toolbox http://cloudimaging.blogspot.com.au/p/about-project.html

NeCTAR Research Cloud...



NeCTAR is funding four programs



NeCTAR – slides by Dr Nigel Ward of NeCTAR

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.



Glue = Galaxy



Tools

<u>Thumbnail</u> identify nuclaei

MILXVIEW TOOLS

Extract 2D slice from an image

Image Registration

- <u>Register using affine transforms</u> on two images
- <u>Register using rigid transforms</u> on two images

Image Segmentation

<u>Segmentation</u> using wm gm and csf

CTE Surface

- <u>Reorient images</u> for CTE surface
- <u>Full CTE Surface pipeline</u> on list of images

CTE

 <u>CTE create raw data</u> on list of images

CT Reconstruction

X-TRACT TOOLS

 <u>CT Reconstruction</u> Create a slice from a sinogram

Cellular Imaging

- <u>Find nuclei</u> identify nuclei
- Find cells from nucleus donut Identify cells given a mask of nuclei
- Find cytoplasm from nucleus donut Identify cytoplasm given a mask of nuclei
- Filter objects by morphology filter objects in an image by their morphological properties
- Find lines Find lines
- <u>Find dots</u> Find dots



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

- 🛱

Extract 2D Slices

Extracts 2D slices in three directions from a 3D image Input image ← 138_T1_MRI.nii.gz

- Galaxy	Analyze Data	Workflow	Shared Data -	Visualization +	Admin	Help -	User-	
Tools Image Processing • Emboss an image • Edges find in image	Extract 2D slice (version 1.0.0) Input Image: 124: 138_T1_MRI.nii.gz Froguta							
<u>Thumbnail</u> identify nuclaei MILXVIEW TOOLS <u>Extract 2D slice</u> from an image Image Registration	This tool extrac page.	ts 2D slices i	n Axial, Saggital a	nd coronal views a	nd inserts	them on to	a html	

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

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



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

Cellular imaging in Galaxy

Tools		History
<u>Get Data</u>		
mage Processing		Unnamed 2.
mage Registration		history
mage Segmentation		
<u>TE Surface</u>		L: ©
<u>TE</u>		8a rqb.tif
T Reconstruction		
ellular Imaging		
Find nuclei identify nuclei		
Find cells from nucleus donut Identify cells given a mask of nuclei		
Find cytoplasm from nucleus		
<u>donut</u> Identify cytoplasm given a mask of nuclei		
Filter objects by morphology		
their morphological		
Find lines Find lines		
Find dots Find dots		
OBJECTS STATISTICS	The second s	
Compute statistics for lines in the image identify nuclei		
EXTRAS		
Overlay image with labelled mask Find lines		
ext Manipulation		
lter and Sort	Constitution of the second	
oin, Subtract and Group		
raph/Display Data		
lorkflows		
IOT KITOWS		
All workflows		· · · · · · · · · · · · · · · · · · ·
All workflows		

Step 1. Find nuclei

🗧 Galaxy	Analyze Data Workflow Shared Data - Help - User -	Using 8.9 Mb
Tools 🌣		History 🌣
Get Data Image Processing Image Registration Image Segmentation	Input image: 2: Nuclei binary mas18a_rgb.tif Nucleoli radius:	Unnamed 2.9 Mb
<u>CTE Surface</u> <u>CTE</u> <u>CT Reconstruction</u> <u>Cellular Imaging</u>	15 Nucleoli radius Largest nuclei radius:	2: Nuclei @ 0 X binary mask for Control EAAT1 n1 Ba rgb.tif
 <u>Find nuclei</u> identify nuclei <u>Find cells from nucleus donut</u> Identify cells given a mask of nuclei 	Largest nuclei radius Smoothing size: 1	<u>Control EAAT1 n1</u> <u>8a rgb.tif</u>
 Find cvtoplasm from nucleus donut Identify cytoplasm given a mask of nuclei Filter objects by morphology filter objects in an image by their morphological properties Find lines Find lines Find dots Find dots OBJECTS STATISTICS Compute statistics for lines in the image identify nuclei 	Smoothing element size Threshold sensitivity: 0.5 Threshold sensitivity Radius of holes: 15 Radius of holes Split objects: Whether to split objects Select image channel: Red channel	
EXTRAS • <u>Overlay image with labelled</u> <u>mask</u> Find lines <u>Text Manipulation</u> <u>Filter and Sort</u> <u>Join, Subtract and Group</u> <u>Graph/Display Data</u> Workflows	Suitable image channel Execute This component finds nuclei in an image of cells	
All workflows		



Output: nuclei binary mask



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

Step 2. Find lines

- Galaxy	Analyze Data Workflow Shared Data + Help + User +	Using 8.9 Mb
Tools 🌣	Find lines (version 1.1.0)	A History
Get Data Image Processing Image Registration Image Segmentation CTE Surface CTE CT Reconstruction Cellular Imaging • Eind nuclei identify nuclei • Find cells from nucleus donut Identify cells given a mask of nuclei • Find colpasm from nucleus donut Identify cytoplasm given a mask of nuclei • Filter objects in an image by their morphological properties • Eind dots Find lots OBJECTS STATISTICS • Compute statistics for lines in	Input image: 1: Control_EAAT1_n18a_rgb.tif Line length: 9 Line length Ignore objects smaller than: 5 Small objects max length Line sensitivity: 2.06 Contrast Smoothing size: 5 Smoothing element size Link Distance: 7 Gap size between links Link quality: 50 Percentage of intensity drop in gaps	Actrocytes 3.0 Mb Actrocytes 3.0 Mb 2: Nuclei Actrocytes 3.0 Mb 2: Attrocytes 3.0 Mb 2: Attrocytes 3.0 Mb 2: Actrocytes 3.0 Mb 2: Attrocytes 3.0 Mb 3: Attrocytes 3.0 Mb 4: Attrocytes 3.0 Mb<
EXTRAS	10.0	Image in tiff format
<u>Overlay image with labelled</u> <u>mask</u> Find lines <u>Text Manipulation</u> <u>Filter and Sort</u>	Line intensity threshold Select image channel: Green channel Suitable image channel	<u>1:</u> ● Ø X Control EAAT1 n1 <u>8a rqb.tif</u>
Join, Subtract and Group Graph/Display Data	Execute	
Workflows All workflows	This component finds lines in an image	- III >



Output: lines binary mask



Last step. Compute statistics on lines per each cell

- Galaxy	Analyze Data Workflow Shared Data - Help - User -	Using 8.9 Mb
Tools 🌣	Compute statistics for lines in the image (version 1.1.0)	History 🌣
Get Data Image Processing Image Registration Image Communication	Input image: 1: Control_EAAT1_n18a_rgb.tif	Image: Optimized systemImage: Optimized systemActrocytes analysis3.0 Mb
Image Segmentation CTE Surface CTE CT Reconstruction Collular Imaging	Lines binary mask: 5: Lines binary mask18a_rgb.tif Cells binary mask: 4: Cells binary mask18a_rgb.tif	<u>6: Lines</u> ● Ø X <u>statistics for</u> <u>Control EAAT1_n1</u> <u>8a_rgb.tif</u>
<u>Find nuclei</u> identify nuclei <u>Find cells from nucleus donut</u> Identify cells given a mask of nuclei	Select image channel: Green channel Suitable image channel	5: Lines
 Find cytoplasm from nucleus donut Identify cytoplasm given a mask of nuclei Filter objects by morphology filter objects in an image by their morphological 	Execute This compute statistics, such as line length, angle, density etc, for lines found in the image.	4: Cells ● ℓ ⊗ binary mask from Filtered objects from Nuclei binary mask for Control EAAT1 n1 8a rgb.tif
properties Find lines Find lines Find dots OBJECTS STATISTICS		3: Filtered ● ℓ ⊗ objects from Nuclei binary mask for Control EAAT1_n1 8a_rgb.tif
<u>Compute statistics for lines in</u> <u>the image</u> identify nuclei <u>EXTRAS</u> Overlay image with labelled		2: Nuclei ● Ø ‰ binary mask for Control EAAT1_n1 8a_rgb.tif
<u>mask</u> Find lines <u>Text Manipulation</u> <u>Filter and Sort</u>		<u>1:</u>
Join, Subtract and Group Graph/Display Data Workflows		
<		

csirc

Output: lines statistics

🗧 Galaxy	_		_	Analyze Da	ata Workflow	v Shared Dal	a- Help-	User⊤	_	Using 8.9 Mb
Tools	\$	number	LineNo	LineMean	LineLength	LineDensity	LineAngle	LineAngleVar	^	History 🌣
Get Data	*	1	7	35.4068	7.57143	0.0155106	77.2432	0.48947		
Image Processing		2	32	43.2456	12.4062	0.0448564	-49.0559	0.72009		
Image Registration		3	63	44.2267	11.1905	0.0873191	-50.9927	0.866407		Actrocytes 3.0 Mb
Image Segmentation		4	40	36.3367	14.3	0.083513	-4.05148	0.722066		anaiysis
CTE Surface		5	22	32.76	9.86364	0.046273	-41.919	0.477815		6: Lines @ / X
CTE		6	57	45.3847	13.5088	0.0297461	-9.71154	0.384416		statistics for
CT Reconstruction		7	81	34.5088	13.4074	0.0372739	7.39296	0.816632		Control EAAT1 n1
<u>CT Reconstruction</u>		8	7	46.692	8.57143	0.00372208	-33.3049	0.60562	Ξ	<u>8a rgb.tif</u>
<u>Cellular Imaging</u>		9	39	31.2907	11.6154	0.0386059	75.6941	0.838289		5 Lines @ 0 %
 Find nuclei identify nuclei 		10	35	44.8682	12.6571	0.0705621	-80.1867	0.714489		binary mask for
 Find cells from nucleus donut 		11	21	33.4321	19.1429	0.016149	17.3238	0.753558		Control EAAT1 n1
Identify cells given a mask of		12	24	49.0239	12.7083	0.0896018	-15.2713	0.503202		<u>8a rgb.tif</u>
nuclei		13	10	31.0749	18.6	0.0458446	56.0637	0.655443		4: Colls @ // \$2
<u>Find cytoplasm from nucleus</u>		14	70	49.5411	14.8429	0.0730073	-21.0695	0.841014		binary mask from
donut Identify cytoplasm		15	26	43.2417	17.0769	0.0975919	-25.5584	0.427264		Filtered objects
given a mask of nuclei		16	55	48.3645	11.7091	0.0875772	-11.9956	0.672437		from Nuclei binary
 Filter objects by morphology 		17	8	40.5262	18.625	0.0952381	-20.7891	0.425022		Mask for Control EAAT1 n1
their morphological	E	18	53	40.6744	12.6226	0.0594804	-59.2297	0.756359		8a rab.tif
properties		19	72	47.3853	14.2361	0.0913362	2.8787	0.450188		
 Find lines Find lines 		20	40	48.2595	14.775	0.108431	3.29259	0.524956		3: Filtered ● Ø 🖇
- <u>rind lines</u> rind lines		21	56	35.3831	14.875	0.0531583	-10.4459	0.885072		objects from Nuclei
 <u>Find dots</u> Find dots 		22	96	33.1553	13.2917	0.0839103	-6.81754	0.492923		Control EAAT1 n1
OBJECTS STATISTICS		23	220	36.53	13.0455	0.0867957	-7.19961	0.818244		8a rgb.tif
Compute statistics for lines in		24	23	45.4204	15.1739	0.106338	-8.94512	0.548168		
the image identify nuclei		25	32	38.7071	14.75	0.0684685	-16.5571	0.804267		2: Nuclei
FYTRAS		26	57	39.5403	13.7018	0.0914344	-1.13796	0.436127		Control EAAT1 n1
EATINAS		27	103	43.5583	11.7087	0.0791557	32.5731	0.759171		8a rgb.tif
 Overlay image with labelled mask Find lines 		28	26	41.9683	20.3077	0.109802	27.7305	0.845859		
mask Find lines		29	79	53.5965	15.0633	0.0738868	-88.1837	0.596194		
Text Manipulation		30	106	40.5542	12.1321	0.0742693	-11.0509	0.693254		Control EAAII n1 8a rob tif
Filter and Sort		31	30	42.3675	12.8333	0.052477	1.93498	0.792295		<u>ou rusa</u>
Join, Subtract and Group		32	20	42.1232	13.05	0.0819763	81.0819	0.516809		
Graph/Display Data		33	137	37.4242	12.0949	0.0682518	-21.1971	0.729491		
Workflows		34	100	40.5995	11.72	0.0946852	1.73017	0.608366		
- All workflows	_	35	34	29.1543	14.0882	0.0623438	65.9247	0.769496		
	111	36	78	40.8803	11.4744	0.0726261	-1.64653	0.844919		
×	111	27	57	41 1741	12 1229	0.0056604	85 0204	0 692720	Ŧ	

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

Complete Astrocytes analysis workflow

Astrocytes analysis workflow can be reused with other image data



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

CSIRO

NeCTAR Imaging Toolkit Production Deploy





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

CSIRO



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

CSS TCP www.csiro.au



HCA-Vision High Level Functionality

ID	FUNCTION	SHORT DESCRIPTION
H.01	Detect nuclei	Detect nuclei in a 2D microscope image
H.02	Detect nuclei from cytoplasm holes	Detect nuclei from absence of stain of cytoplasm
H.03	Detect cells with nuclei	Detect cells using nucleus image as a mask
H.04	Detect cells without nuclei	Detect cells without using nucleus image as a mask
H.05	Detect neurons with nuclei	Detect neurons from a neurite outgrowth image using nucleus image as a mask
H.06	Detect neurons without nuclei	Detect neurons from a neurite outgrowth image without using nucleus image as a mask
H.07	De-clump touching objects	Separate any touching objects in an image, such as touching nuclei or cells
H.08	Label objects	Label objects such nuclei or cells in a binary image
H.09	Get object stats	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.
H.10	Detect cell from nucleus donuts	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.
H.11	Detect dots	Detect dots in a 2D image or a cell
H.12	Detect lines	Detect line structures in a 2D image or a cell
H.13	Get dot stats	Retrieve statistical features of the detected dots, including area, perimeter etc.
H.14	Get line stats	Retrieve the statistical features of detected line structures
H.15	Cell Scoring	Count negative and positive cells, measure integrated and average intensity of negative and positive cells.

IDs H.xx – Application area: cell features of microscopy images

X-TRACT High Level Functionality

ID	FUNCTION	SHORT DESCRIPTION
XP.01	Sinogram creation	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).
XP.02	Ring artefact removal	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.
XP.03	Dark current subtraction	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.
XP.04	Flat field correction	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.
XP.05	Positional drift correction	The function is used for correction of transverse drift between related experimental images. Image drift is assessed by cross-correlating pairs of images.
XP.06	Data normalisation	Data normalisation Including normalisation to a user-defined region
XP.07	TIE-based phase extraction	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).
XCT.01	FBP CT reconstruction	Filtered back-projection (FBP) parallel-beam CT reconstruction
XCT.02	FDK CT reconstruction	Feldkamp-Davis-Kress (FDK) cone-beam CT reconstruction
ХСТ.03	Centre of rotation	Automated calculation of the centre of sample rotation in a CT scan from experimental X-ray projections, sinograms or reconstructed axial slices.
XCT.04	CT Reconstruction Filters	The choice of available CT reconstruction filters will include at least the Liner-Ramp, Shepp-Logan, Cosine, Hamming and Hann filters.
XCT.05	ROI reconstruction	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.
IDs XP.xx	Application are	a: data processing functions

IDs XCP.xx – Application area: CT reconstruction functions

MILXView High Level Functionality

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

IDs MC.xx – Application area: neuro-imaging analysis, cortical thickness estimation (CTE)

IDs MS.xx – Application area: neuro-imaging analysis, CTE surface

IDs MP.xx – Application area: neuro PET analysis, PET PVC

IDs MR.xx – Application area: neuro PET analysis, PET SUVR