Solving Hierarchal Neuroscience Problems With Parsl Matthew Madany, <u>madany@ucsd.edu</u>

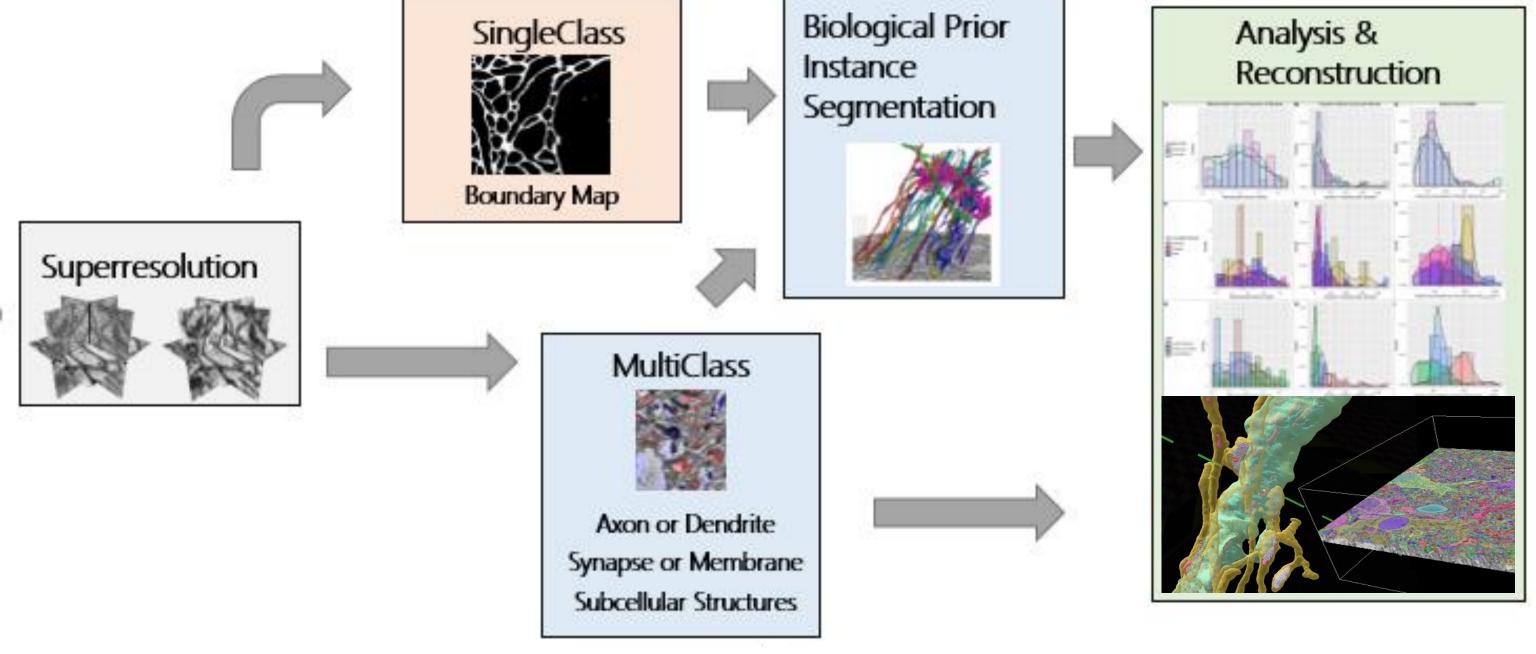




National Center for Microscopy and Imaging Research

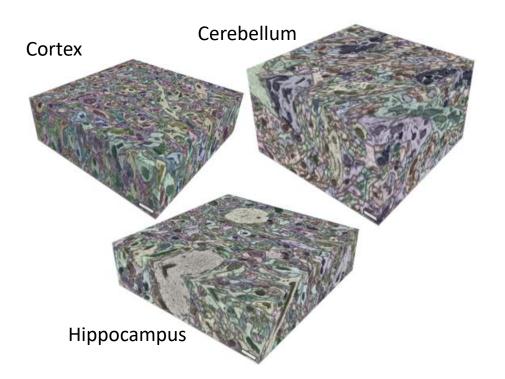


SIDE SAN DIEGO SUPERCOMPUTER CENTER



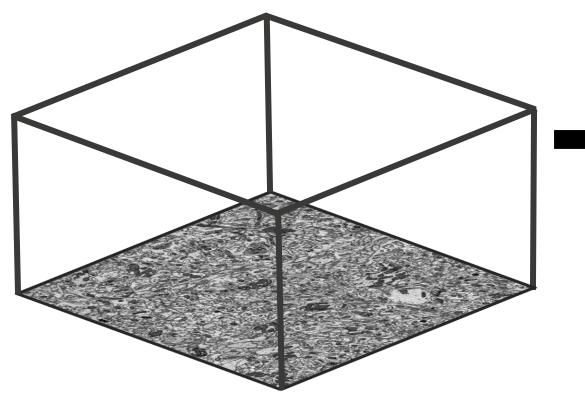
Volume Electron Microscopy A.I. Applications

Labeled Volume Data from Brain Images

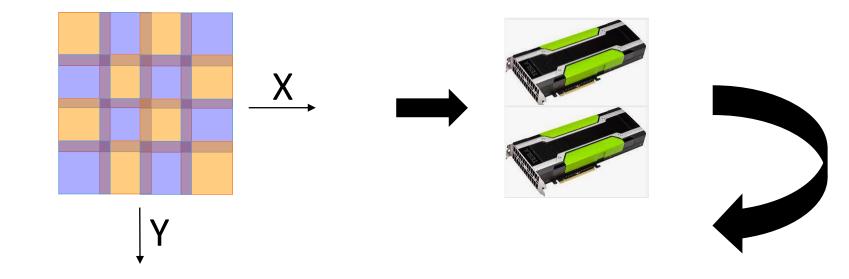


Challenges in Application Design

- Defining A parallelization scheme For efficient GPU and CPU utility
- Defining a data structure scheme that is scalable and portable



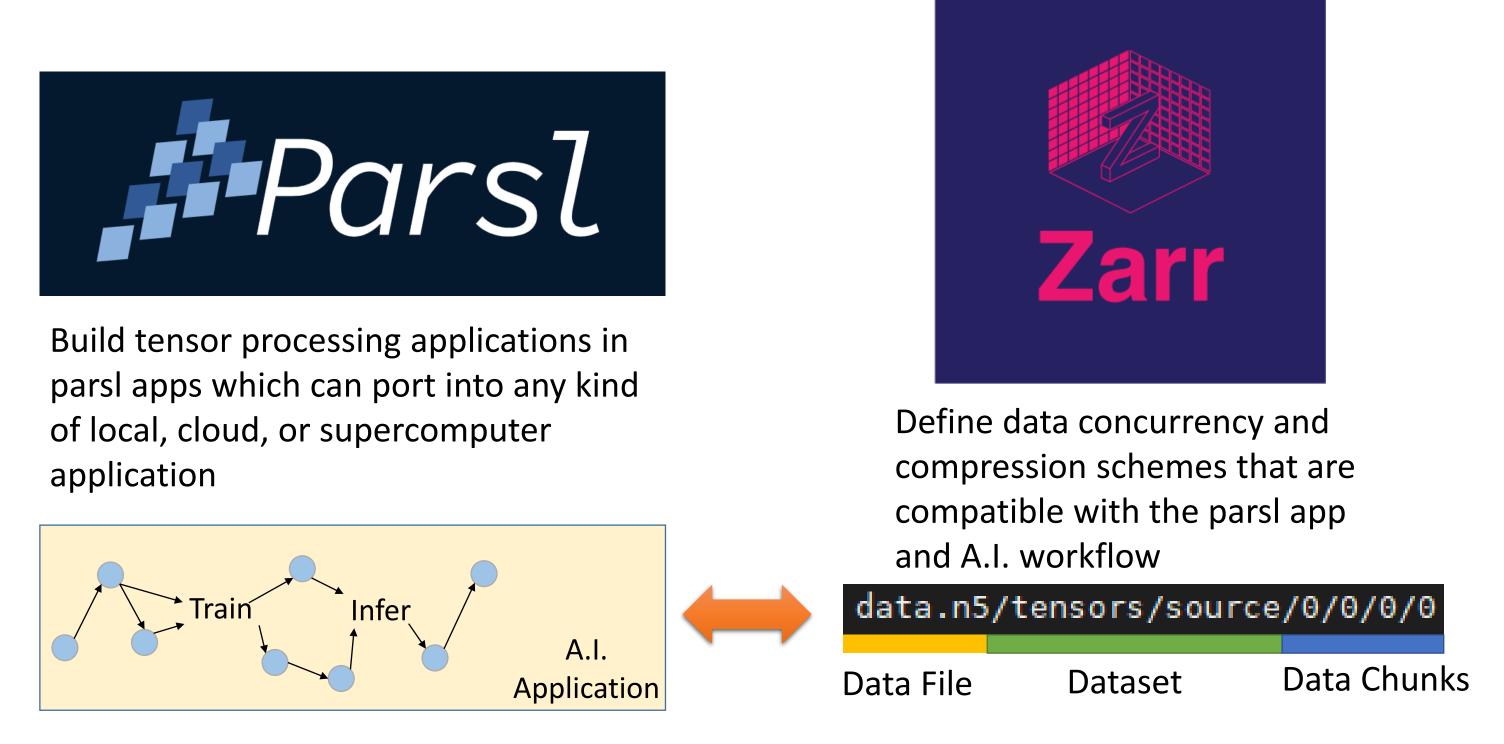
Factor 3D volume into 3D windows



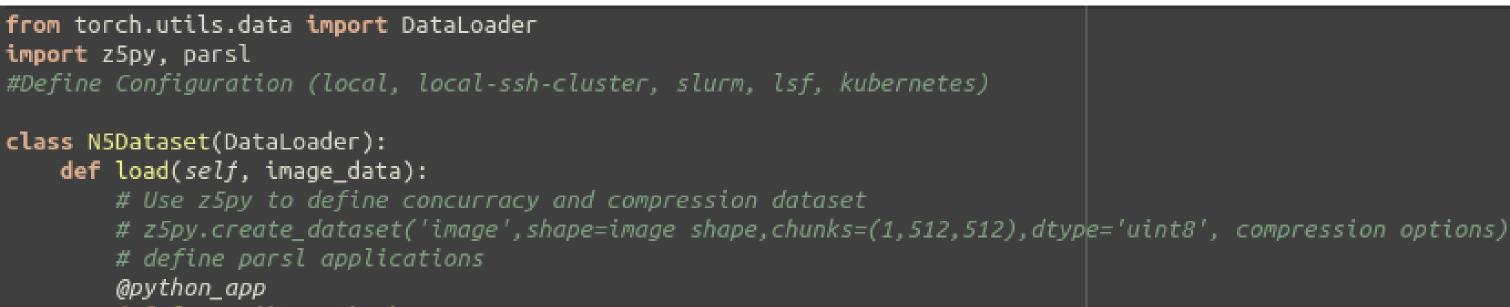
Design Criterion:

3D volume has to be factored into windows (tensors) compatible to GPU specs, while this process must also compress and store data compatible to storage location, (balancing shared filesystem capacity, node storage capacity, compression, inodes)

Scalable Volume ML/GPU Applications must therefore define: Workers to import and format data, workers to port this data through GPUs, workers to compile outputs back to original volume while dealing with complex multidimensional processing pipelines that wrap around simplified machine-learning experiments. Solution: Combine a parallel execution engine with a concurrent data structure scheme



Building Scalable volumetric A.I. applications becomes much easier



def factor(iteration):
 # Large image data -> tensor field -> single tensor

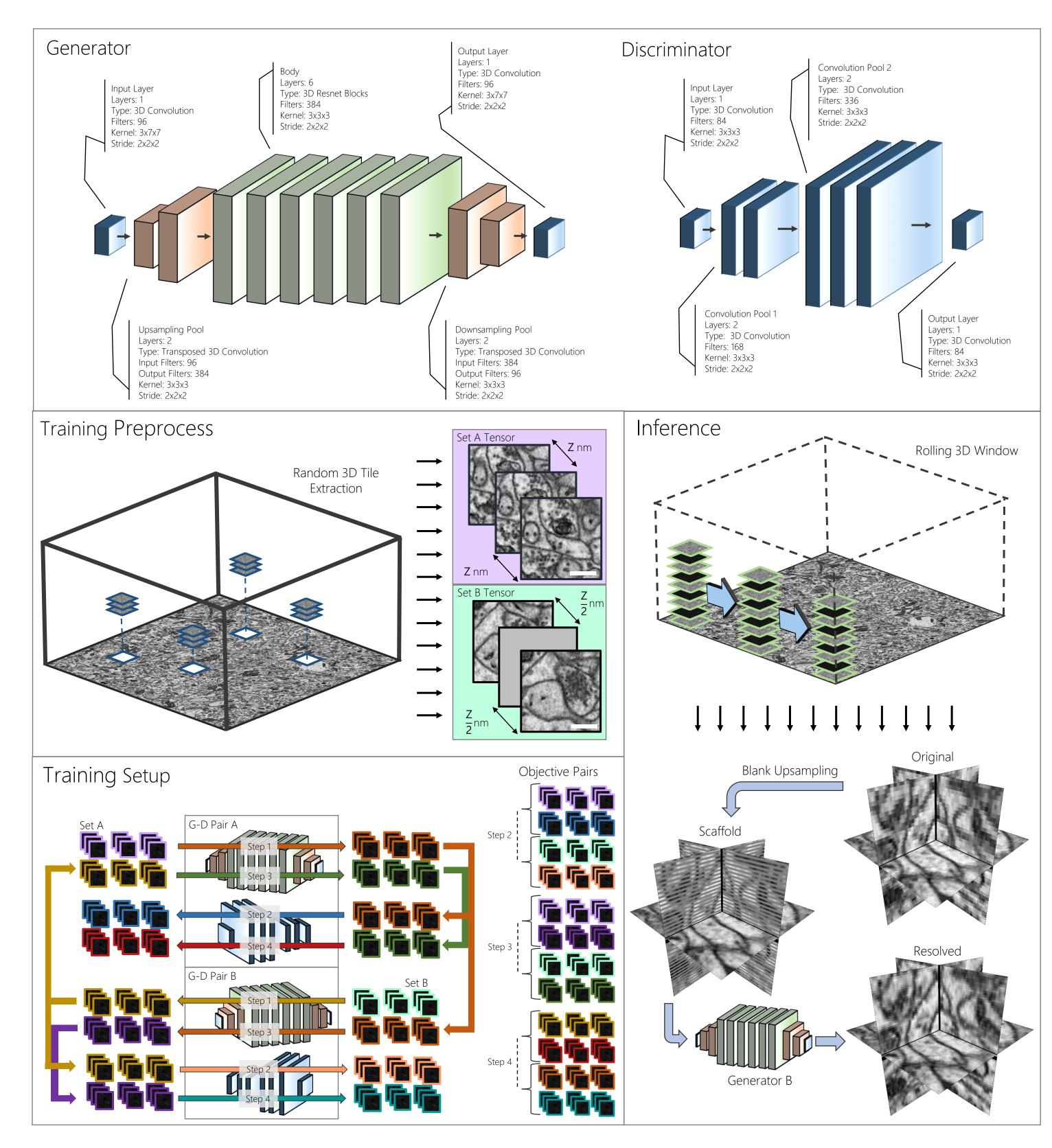
Data achieves very high dimensionalities and efficiencies Dataset x Batch x Channels x Z-Depth x X-Dimension x Y-Dimension x Tensor-iteration

Configuration options passed to Parsl and z5py will naturally load-balance these needs:

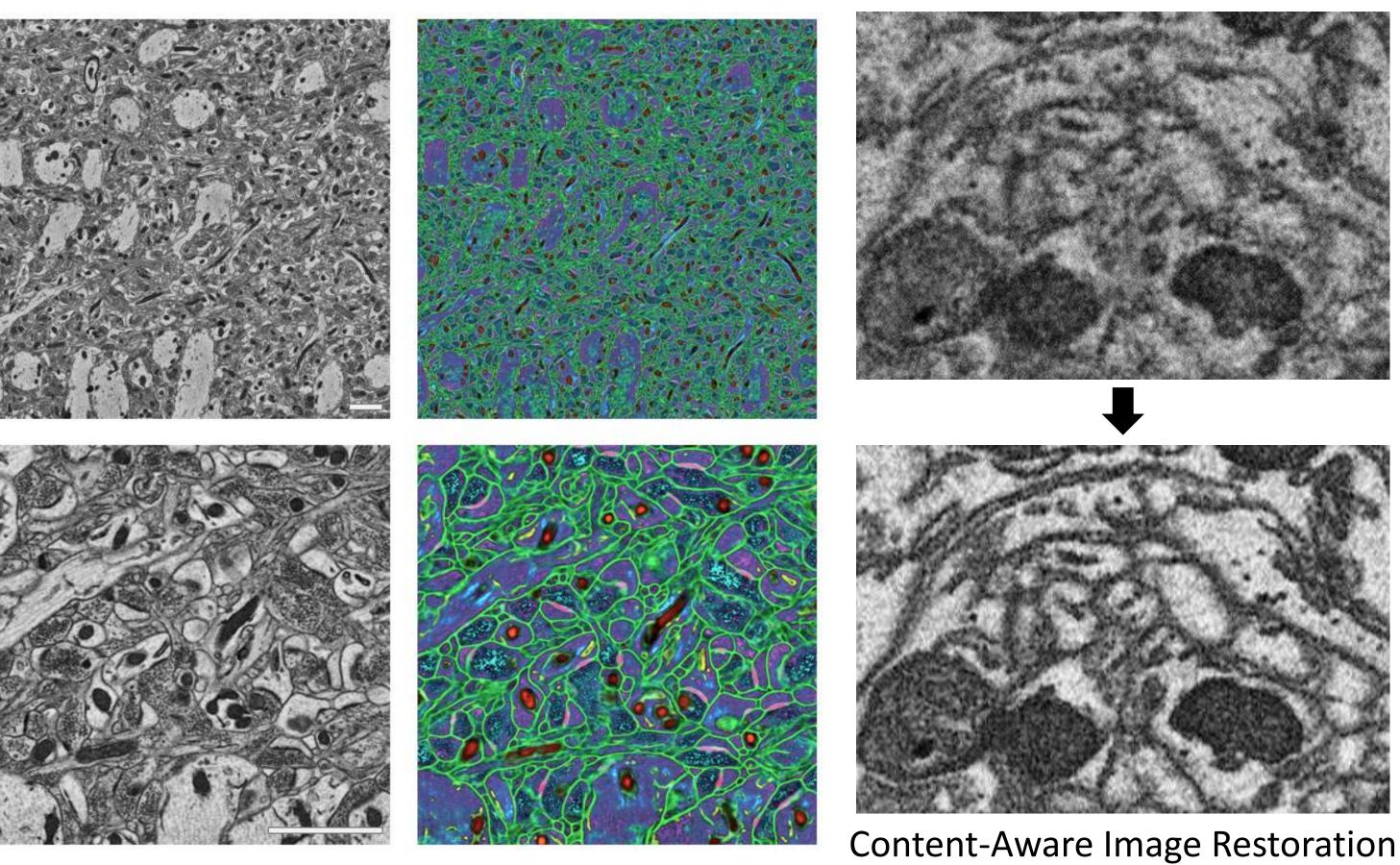
- Efficiency in GPU/CPU processing
- Wrangling large data into deep learning frameworks
- Easy design of high-dimensionality processing pipelines
- Data compression cost and benefit
- Porting of applications to heterogonous node definitions

Altogether enable the testing of applications at scale that we wouldn't normally even consider building in the first-place due to parallelization challenges.

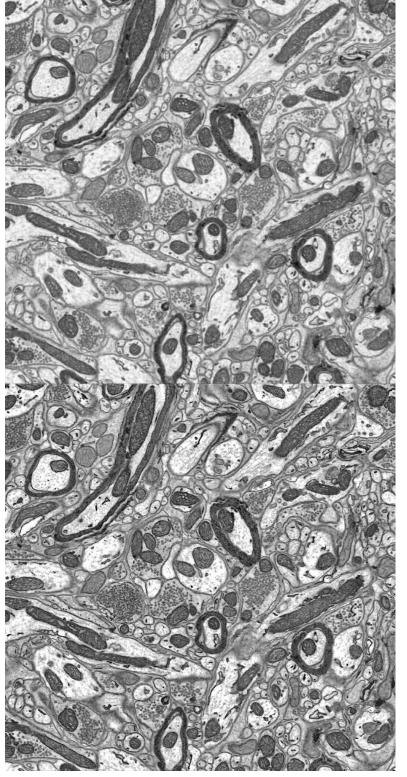
Electron Microscopy Superresolution

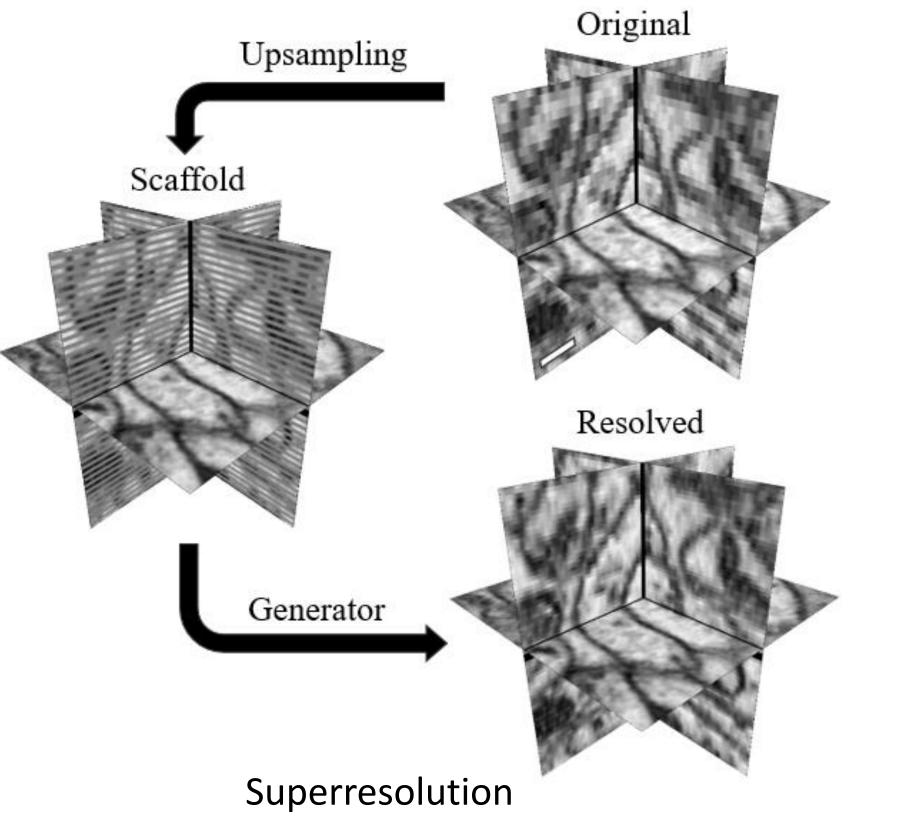


Examples of Applications



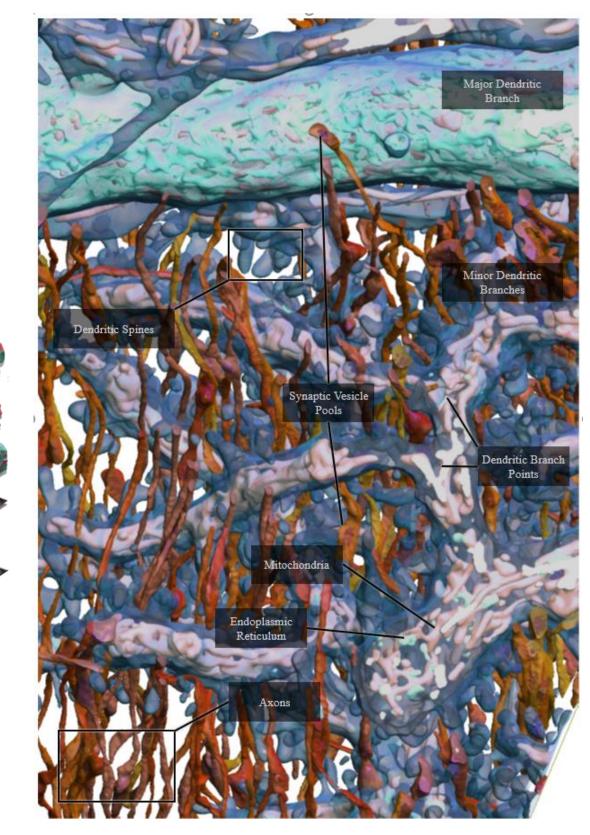
MultiLabel Voxel Classification

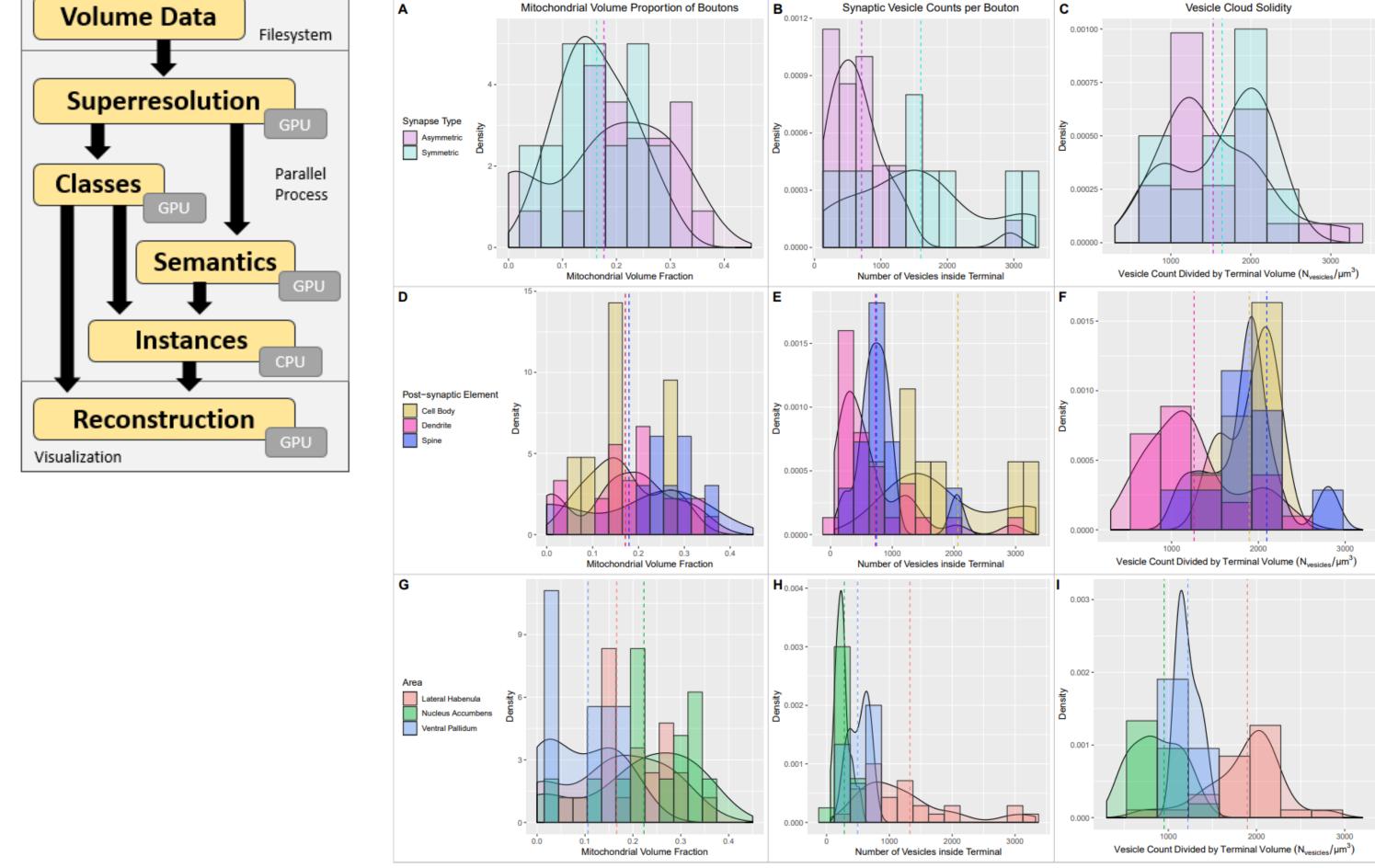


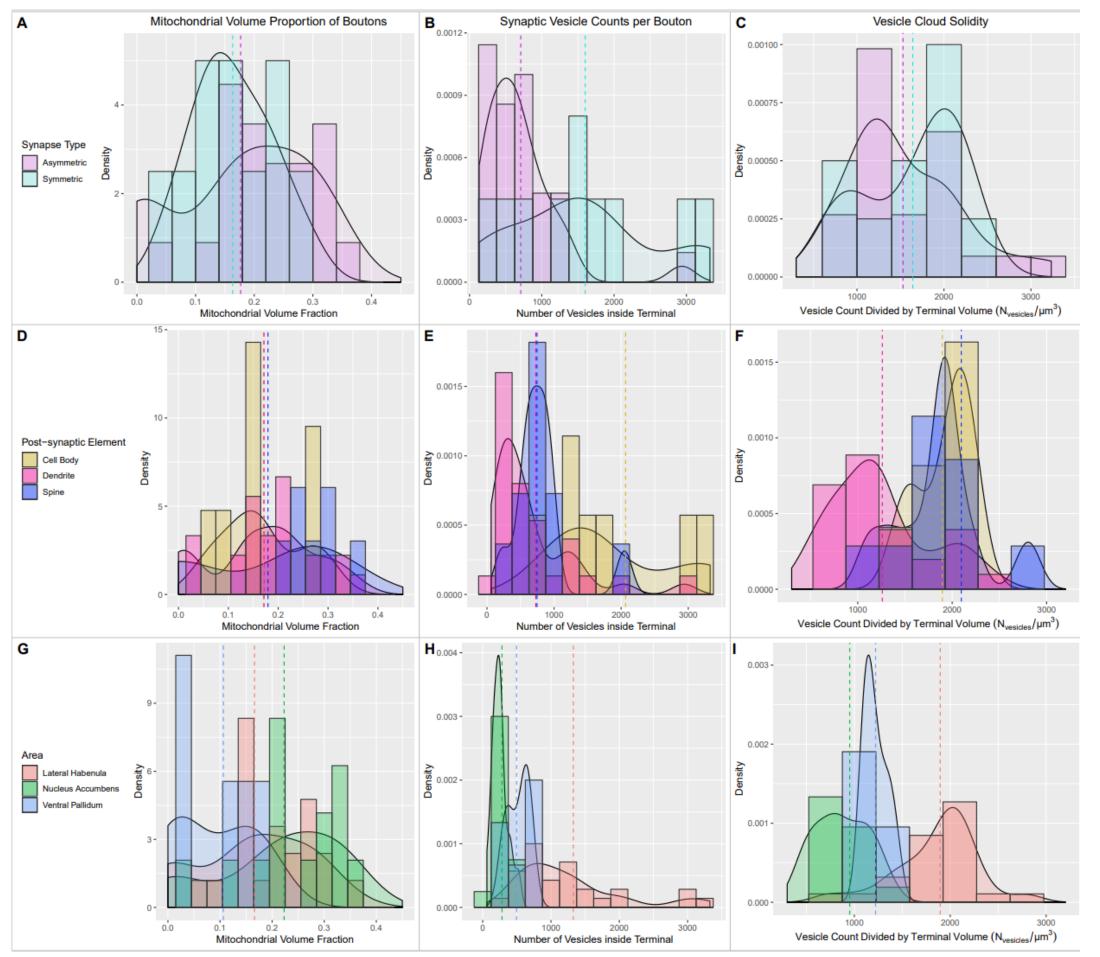


Microscopy Domain Translation

Examples of Reconstructions, Analyses, and Composable Workflows







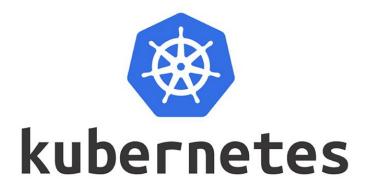
Performance

CPU / Mem

						· · ·	
器 Kubernetes / Compute Resources / Names	pace (Pods) ං				₽ < 02	2020-08-14 20:02:40 to 2020-08-17 00:03:11 UTC 🗸 👌 😋	
datasource default - namespace ncmir-mm -							
~ Headlines							
CPU Utilisation (from requests) CPU Utilisation (from limits)			Memory Utilization (from requests) Memory Utilisation (from limits)				
20.08%		2.748%		0.6386%		0.1868%	
~ CPU Usage							
			CPU Usage				
— cil-cdeep3m-b-8813 — cil-cdeep3m-c-8817 — cil-cdeep3m		f-bnv92 — jupyter-madany-40ucsd-2eedu –	– proxy-697946968b-lzwk5 – tiller-deploy-6585d5ffdf-wl6ms –	0 08/16 00:00 08/16 02:00 08/16 04:00 08/16 06:00 volbatch-1-2nts8 – volbatch-1-5b9k9 – volbatch-1-jkwb7 – vo ch-11-416/r – volbatch-11-9k/52 – volbatch-11-nxhnb – volbatch-	ibatch-1-vn4bn — volbatch-10-hft8g — volbatch-10-nztsd —	volbatch-10-t44xd — volbatch-10-xkfbh — volbatch-100-bhppn	
– volbatch-115-8kkdt – volbatch-115-xdsb8 – volbatch-116 ~ CPU Quota	-dqpz6 — volbatch-117-4dhr8 — volbatch-118-4qc	qvn — volbatch-119-kktsc — volbatch-12-g	jmccs — volbatch-12-jr55n — volbatch-12-ngwwv — volbatci	h-12-v8kmq volbatch-120-gqqt9 volbatch-121-b7sms volb	iatch-122-r7jfr — volbatch-123-rzdks — volbatch-124-kfzjq —	volbatch-125-vgnn7 — volbatch-126-brtwf — volbatch-127-9x4zq	
			CPU Quota				
Pod volbatch-51-ttlk2		CPU Usage	CPU Requests	CPU Requests %	CPU Limits 12.00	CPU Limits %	
volbatch-195-rn2kj		1.00	1.00	99.97%	12.00	8.33%	
tiller-deploy-6585d5ffdf-wl6ms		0.00	1.00	0.01%	2.00	0.00%	
hub-67cfcd595f-bnv92		0.04	2.00	1.98%	8.00	0.49%	
globus-connect-fs-0							
nrayu-60701/6068h-lzwk5		0.00	0.20	2 17%	4 00	በ 17%	
~ Memory Usage							
			Memory Usage (w/	o cache)			
559 GiB							
373 GIB	Ach		A.M. A.N.				
186 GiB					Man and the second states and the second states		
0 B —	08/15 04:00 08/15 06:00 08/15 08:00	08/15 10:00 08/15 12:00 08/15 14:00	0 08/1516:00 08/1518:00 08/1520:00 08/1522:0	0 08/16 00:00 08/16 02:00 08/16 04:00 08/16 06:00	08/16 08:00 08/16 10:00 08/16 12:00 08/16 14:00	0 08/16 16:00 08/16 18:00 08/16 20:00 08/16 22:00 08/	
— volbatch-1-vn4bn — volbatch-10-hft8g — volbatch-10-nzts	sd — volbatch-10-t44xd — volbatch-10-xkfbh — v	volbatch-100-bhppn — volbatch-101-6jvpw	- volbatch-102-dircg - volbatch-103-r9545 - volbatch-104-c	bbus-connect-fs-0 — hub-67cfcd595f-bnv92 — jupyter-madany-40uc p6lw — volbatch-105-6xs75 — volbatch-106-tf778 — volbatch-10 118-4qqvn — volbatch-119-kktsc — volbatch-12-gmccs — volbat)7-q9qks — volbatch-108-ktgrz — volbatch-109-ds9pc — voll	fdf-wl6ms – volbatch-1-2nts8 – volbatch-1-5b9k9 – volbatch-1-jkwb7 batch-11-4l6ir – volbatch-11-9vk5z – volbatch-11-nxhnb – volbatch-11 volbatch-120-gqqt9 – volbatch-121-b7sms – volbatch-122-r7jfr	
~ Memory Quota							
			Memory Quo	la			
Pod	Memory Usage	Memory Requests	Memory Requests % Memory Limits	Memory Limits % Memo	ory Usage (RSS) Memory Usage (C	ache) Memory Usage (Swap)	
cil-cdeep3m-d-8793		8.00 GiB	- 24.00 GiB				
cil-cdeep3m-c-8804		2.00 GiB	- 8.00 GiB				
cil-cdeep3m-a-8809		8.00 GiB	- 24.00 GiB				

GPU





PRP PACIFIC RESEARCH PLATFORM





National Center for Microscopy and Imaging Research



SIDE SAN DIEGO SUPERCOMPUTER CENTER