LIMPID/BisQue

A scalable infrastructure for reproducible, image driven data science

Principal Investigators

B. S. Manjunath, ECE, UCSB
Tresa Pollock, Materials, UCSB
Robert Miller, Marine Sciences, UCSB
Nirav Merchant, CyVerse, U. Arizona
Amit Roy-Chowdhury, ECE, UC Riverside

BisQue

Dmitry Fedorov
Kristian Kvilekval
Christian Lang

Applications

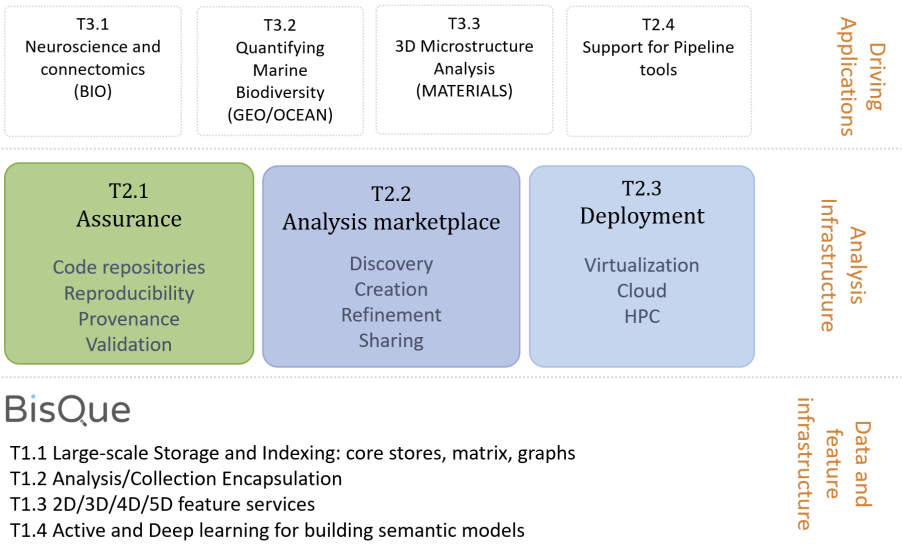
Mclean Echlin, Materials
Marat Latypov, Materials
Anandasankar Ray, Neuroscience
William Smith, Neuroscience

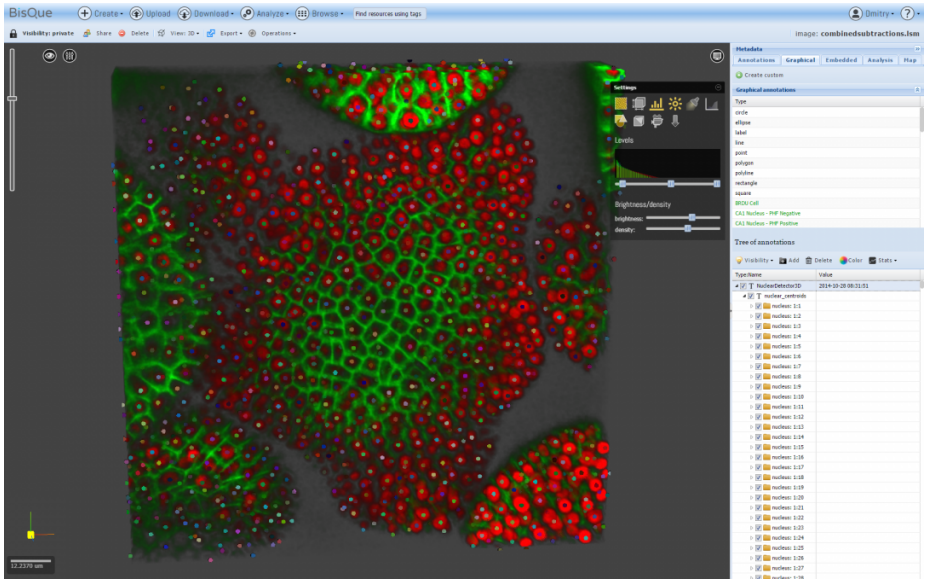
NSF #1664172 (EAGER, 2016-18)
NSF #1650972 (SSI, 2017-22)

LIMPID/BisQue Overview

- LIMPID is built on cloud-based analysis platform BisQue
- Management, analysis, and sharing of images and metadata for large-scale data science
- Flexible and scalable query system across network of multimodal data items
- Module system for scalable integration of analysis tasks over images and metadata
- 200+ life science image and video formats
- Analysis marketplace: easy sharing and discovery of analysis modules

- *More information on BisQue:* http://bioimage.ucsb.edu/bisque

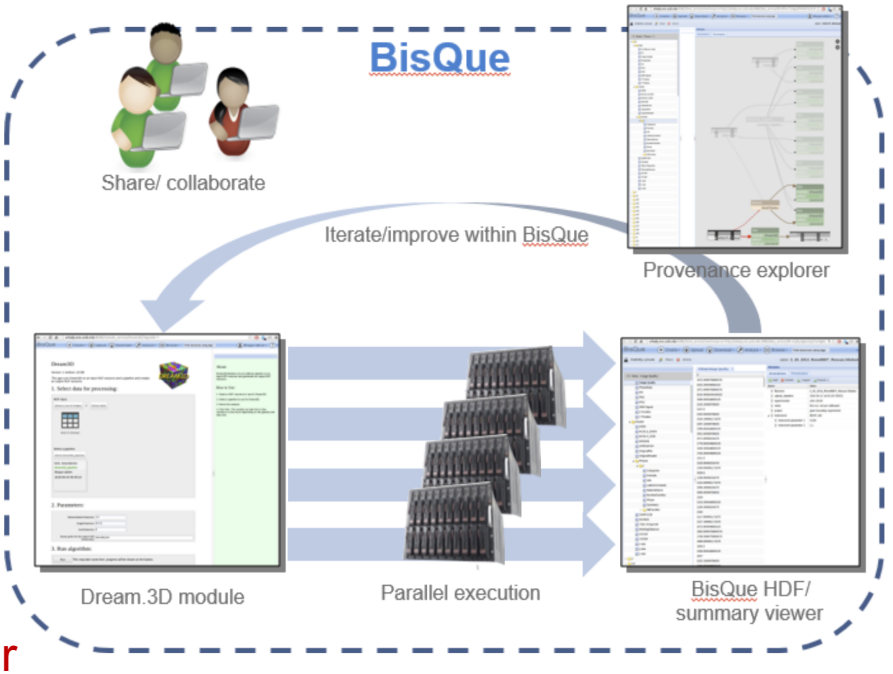


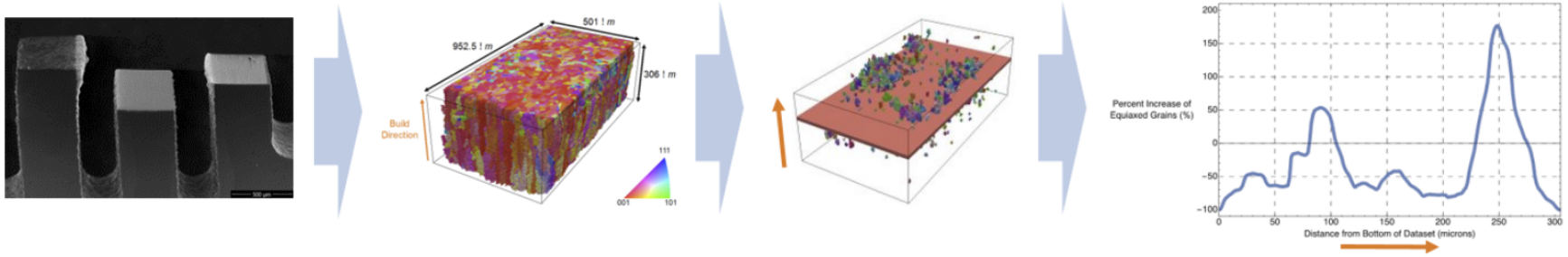


Dream.3D Integration

(Use Case: Materials Structure Analytics)

- Large acquired Materials Science datasets analyzed in desktop apps such as Dream.3D
- Dream.3D pipelines and execution integrated in BisQue
- Datasets and pipelines can be shared with collaborators
- Parameter variations are parallelized over compute grid
- Large datasets viewable in web browser
- Provenance tracking for improved repeatability

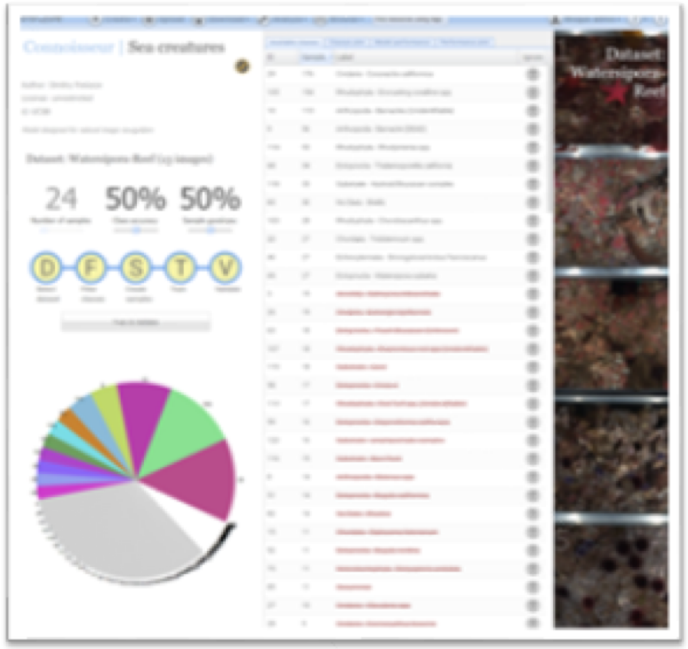


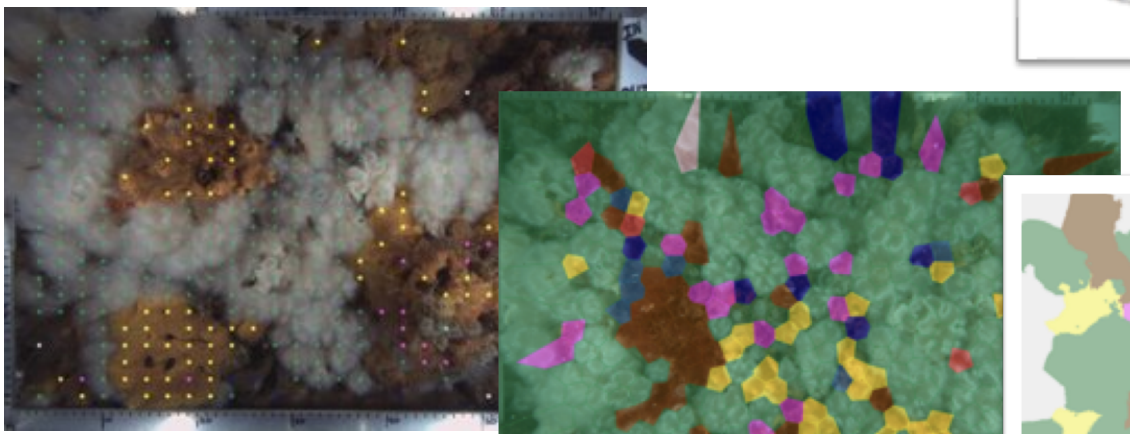


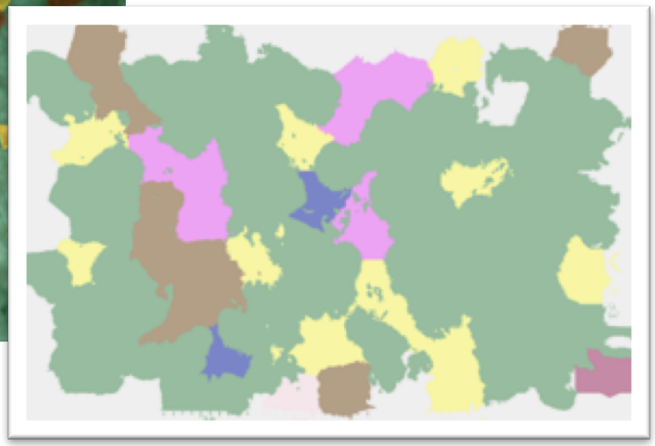
Deep Learning Integration

(Use Case: Quantifying Marine Biodiversity)

- **Goal:** Automate annotations without engineer in reduced time for specific datasets
- Metadata driven training set selection
- Automatic parallelization
- Integrated validation (model selection)
- Multiple imaging modalities: photographs, satellites, microscopes



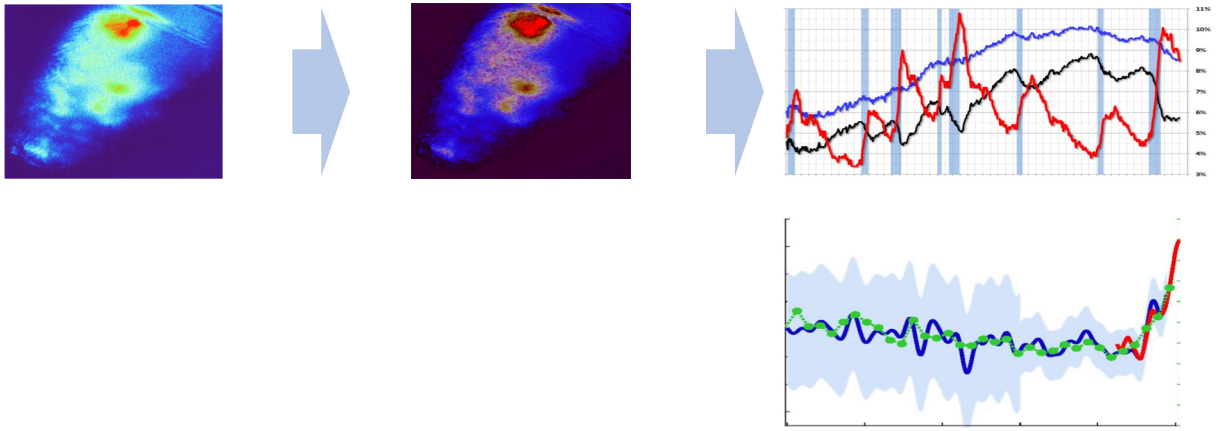




Deep learning method development

(Use Case: live Calcium imaging in neurons)

Application: Quantify and map changes in fluorescent signal in 4D data acquired from live imaging of neuronal activity in insect nervous system



- Requires registration of images to compensate for movement.
- Automated identification of neurons from background based on shape, position and fluorescent changes.
- In a post processing phase quantify fluorescent changes over time for each neuron identified.
- Create plots of runs for evaluation as supervised parameters are altered, to create summarized data.