

# From Images to Knowledge with ImageJ & Friends

virtual conference

**Nov 30 - Dec 2, 2020**

Stephan Preibisch, Stephan Saalfeld, Anna Kreshuk,  
Pavel Tomancak and Virginie Uhlmann

hhmi | janelia  
Conferences



## DeepCell: Deep learning for single cell image analysis

**Tutors:** David Van Valen (vanvalen@caltech.edu)

**Session 1:** 2020-12-02 00:00 UTC – 2020-12-02 04:00 UTC

**Session 2:** 2020-12-02 17:00 UTC – 2020-12-02 21:00 UTC

# DeepCell: Deep learning for single cell image analysis

David Van Valen MD, PhD

California Institute of Technology

## Abstract

Deep learning is changing the relationship between biologists and imaging data by significantly reducing the difficulty of extracting quantitative data from images. For the past two and a half years, my lab has been working to adapt these methods to single cell image analysis. Rather than solely focus on algorithms, our approach has been to co-develop training datasets, models, and deployment software. The results of these efforts are reflected in DeepCell, our lab's collection of software tools for performing each of these three tasks, as well as training data and models that we have developed. In this workshop, we propose to show participants how to perform single cell analysis in tissues by accessing existing models and developing new ones using the tools available in DeepCell.

## Proposed Workshop

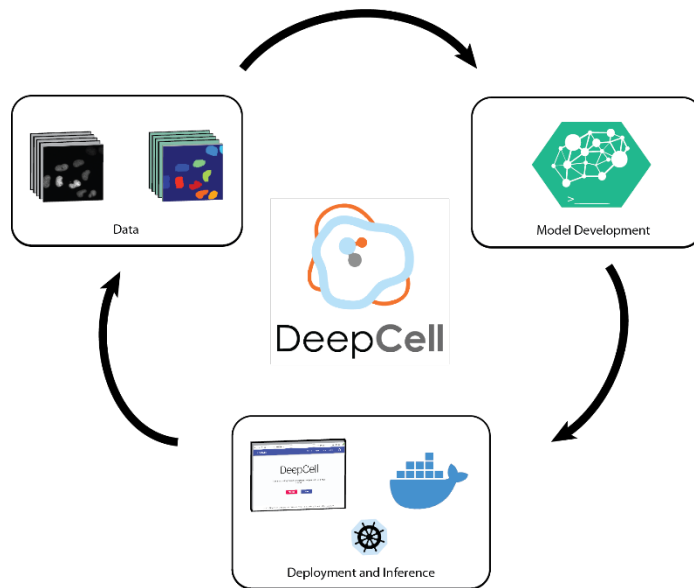


Figure 1: Successful application of deep learning to problems in cellular image analysis requires co-development of training datasets, models, and deployment software.

in laboratory settings. A more detailed description of our software tools can be found at <https://www.deepcell.org/about>.

Our proposed workshop will have two practical parts. The first part will focus on our collection of pre-trained models and be targeted towards users with less experience in programming and image analysis. We currently have two collections of pre-trained models. One is geared towards live cell imaging – these models can do nuclear segmentation and whole cell segmentation in static frames, and nuclear tracking/lineage construction in movies. The other set is geared towards multiplexed images of tissues. These models can perform nuclear and whole cell segmentation and tissues. These two collections of models are powered by expansive training data (~400,000 and ~1,300,000 cells

Getting deep learning to work in practice requires much more than just deep learning, as outlined in Figure 1. Diverse and accurately labeled training data are key to developing performant deep learning models, and robust, easy-to-use deployment software is necessary if these models are to see widespread use. For the past two and a half years, my research group has been working on solving these challenges for single cell image analysis. Our primary focus has been on creating general purpose deep learning models for analyzing live-cell imaging movies and multiplexed images of tissues. These efforts have led us to annotate a diverse set of images across a variety of image types, develop new deep learning approaches to single cell analysis, and created cloud-native software for deploying deep learning-enabled image analysis pipelines

for each model collection respectively) and can be used out of the box on reasonably similar data. These models are hosted on <https://www.deepcell.org>, and can be used with either the web portal, and ImageJ plugin, a command line tool, or locally through a docker container. This part of the workshop will show users how to perform inference using each of these different approaches.

The second part of the workshop will focus on developing models for single cell analysis. Developing new models has two components – creating training data and training a model. We will show users how to use Caliban, our lab's software for annotating higher-dimensional biological images (multiple channels, 3 spatial dimensions, and 2 spatial/1 temporal dimension). Caliban currently supports creating annotations from scratch, as well as editing existing annotations (an essential part of creating large datasets through a human-in-the-loop approach). We will also show users how to train a model in Google Colab using our lab's deep learning library (<https://www.github.com/vanvalenlab/deepcell-tf>), and how to benchmark models using our metrics package (<https://www.github.com/vanvalenlab/deepcell-toolbox>).

The workshop will be organized by myself, our lab's software team (Will Graf, Tom Dougherty, Geneva Miller), and the postdocs/graduate students who have been spearheading this effort (Erick Moen, Noah Greenwald, Morgan Schwartz). I will make myself available for the entire day for the day of the tutorial.