ELEC576 Project Proposal

Segmentation of Yeast Fluorescence Microscopy Images using Deep Neural Networks

Jihwan Lee, Zhuohe Liu, Xiaoyu Helen Lu

Nov. 7, 2018

Introduction and Background

Fluorescence microscopy is a powerful tool to visualize cellular and subcellular structures and processes. Fluorescent proteins are a preferred tool of imaging because they have genetically encoded fluorophores and hence can be non-invasively introduced in the form of the corresponding gene. A fluorescent protein can be fused to virtually any protein (Enterina et al. 2015). In addition, the most common forms of fluorescent proteins that are derived from jellyfish and corals do not require cofactors as the fluorophores of these proteins form in autocatalytic manner in the presence of oxygen (Reid and Flynn 1997; Subach and Verkhusha 2012). Our lab proposes a ultra-high-throughput way of screening brighter fluorescent proteins (FP) under the microscope at a single-cell level at which individual yeast cells contain a single FP variant. To analyze fluorescence intensity at a single-cell level, a process called segmentation is necessary. Segmentation is a step that allows a computer to differentiate regions that represents cells from the background. Because fluorescence images are noisy, simple thresholding often fails and therefore, we plan to investigate a machine learning approach of cell segmentation.

Goal

The goal of this project is to adapt existing deep neural architectures to the application of yeast image segmentation. We anticipate to accomplish this goal in a succession of milestones. First, we will adapt deep networks, such as UNet and SegNet, to perform binary classification of the image. Afterwards, we plan to train deep nets to identify individual cell boundaries in cropped images (32x32) based on previous masking. Once the net is proven working in cropped images, we will adapt the net back to the original 2048x2048 image to improve the throughput and fit the actual data we acquired from experiments.

Feasibility

Algorithm feasibility: There had been proved that Deep Convolutional Neural Network (DCNN) is powerful in doing cell segmentation (Sadanandan et al, 2017, and the algorithm had been implemented in software CellProfiler), so we should be able to achieve decent cell segmentation using DCNN by making it fully adapted to our dataset.

Data feasibility: We have already acquired thousands of images that we can train and test on. To avoid manually handling labels for all these images, we have already made a graphics generator that can generate fake training data with labels.

Computational power feasibility: While the original dataset is too big to directly work with, we can perform data preprocessing to crop the large images into smaller ones since the objectives (i.e., single cells) we are trying to recognize is rather small. Preliminary training will be perform on cropped images and can be finished on laptops or Rice NOTS.

Method

There are two challenges of this project considering the time and data resources we have. First, unlike everyday photos, fluorescence images can be much noisier due to the low emission light intensity of the biological samples, and have a higher dynamic range (Fig. 1). Therefore, a true label of the pixel is hard to obtain even using manual approach. In addition, because of the high-throughput nature, resolution of the yeast is limited, whereas the size of the image is maximized (Fig. 1). This means our solution should be compatible with the large size of the images that we capture. Besides, we want to avoid adding additional channels dedicated for segmentation (such as phase contrast) and only the fluorescence channel(s) available to us.

To overcome the challenges, we plan to code a simple graphics generator for synthetic yeast fluorescence images, and find a way to adapt the network to real images that are thousands times more pixels than smaller training images which we can easily handle. We will explore different model parameters and hyperparameters such as network depth, size of convolutional layers, and optimize the results.

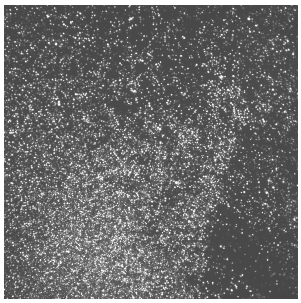

Figure 1. A yeast fluorescence microscopy image. *Fluorescence microscopy image captured by a sCMOS camera has a size of 2048 by 2048, and a resolution of 0.32 μm/pixel, which corresponds to around 20 pixel for the diameter of the yeast cell. Only one color channel is available to use for segmentation.*

Initial Results

We implemented the networks in MATLAB as a proof of concept. The training using 4000 32x32 images, which corresponds to the area of a single full size microscopy image we have, took moderate amount of time to finish. The training accuracy could be as high as 98%, but we have not overfit the model yet. Our initial test of the network on cropped real images yielded reasonable labels (Fig. 3).

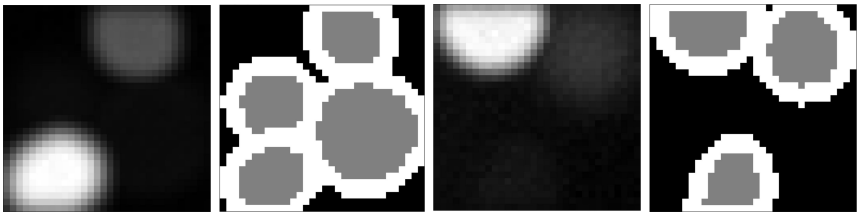

Figure 2. Synthetic images and labels for training. *Synthetic images of yeast and labels can be generated with random numbers drawn from a predefined range of values. Properties include size, orientation, brightness, focus, etc. are modeled in order to approximate the real image. Two pairs of example images and labels (border in white, cell body in gray, and background in black) are shown here. Not all cells are visible because the high dynamic range and low contrast.*

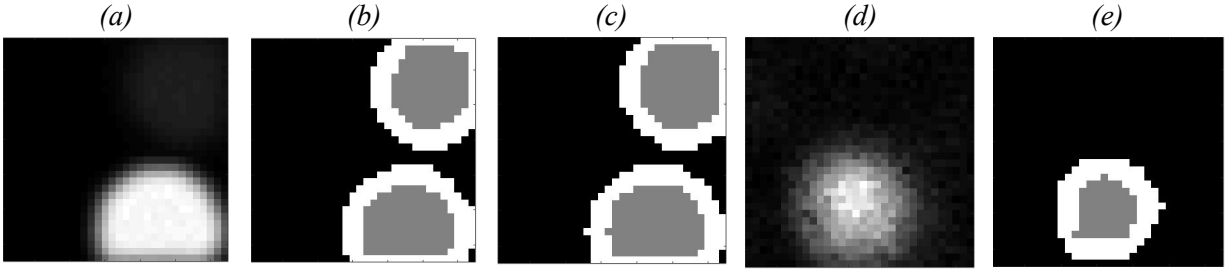

Figure 3. Initial segmentation results. *(a) Sample testing image. (b) Generated testing label from (a). (c) Ground truth label of (a). (d) Cropped real fluorescence microscopy image for testing. (e) The label generated by the network from the real image (d).*

Reference

Enterina, J.R., Wu, L. & Campbell, R.E., 2015. Emerging fluorescent protein technologies. *Current Opinion in Chemical Biology*, 27, pp.10–17.

Reid, B.G. & Flynn, G.C., 1997. Chromophore formation in green fluorescent protein. *Biochemistry*, 36(22), pp.6786–6791.

Subach, F.V. & Verkhusha, V.V., 2012. Chromophore transformations in red fluorescent proteins. *Chemical Reviews*, 112(7), pp.4308–4327.

Sadanandan, S.K., Ranefall, P., Le Guyader, S. and Wählby, C., 2017. Automated training of deep convolutional neural networks for cell segmentation. Scientific reports, 7(1), pp.7860.