Computational Biology

Automated Particle Picking in Cryo-EM

Qinyao He, Dayu Yue, Xi Ye

# Overview

# Our Approach

## Preprocessing

In order to provide the training data for neural network, we cropped positive and negative sample with size 180\*180 from the original large image, the crop was then resized to 64\*64.

In this way we gain 70000 train image with half positive and negative examples.

## Multi-Layer Perceptron

For simplicity and faster training for validation continuous process, a simple multilayer neural network with two hidden layer is built. We treat the input as dimention vector, with 256 output unit in each hidden layer, and the last layer output one value as the probability for the positive [hypothesis](http://dict.youdao.com/w/hypothesis/#keyfrom=E2Ctranslation) (is a particle). And we use binary cross entropy as loss function.

It turns out to be a rather simple model, but we get a validation accuracy of 0.9. Detailed learning process will be covered in next section.

## Convolutional Neural Network

Convolutional Neural Network has recently been proved to have state-of-art performance in image-related tasks. Here we can regard the Cryo-EM as a large image, and we are going to learn and detect some pattern which correspond to the particle. CNN seems to be very good at extracting high level feature from the original image (without preprocessing and feature engineering) automatically.

We follow the practice of VGG-Net, use all 3\*3 convolution filter size, with padding 1 to make sure the size doesn’t change after convolution. And a 2\*2 max pooing is followed every one or two convolution layer. Channel size doubled every one or two max pooing. And two dense layer is added after last convolution and before loss layer.

Generally we stop add more convolution when image size is 2\*2.

## Particle Detection

# Experiments

# Conclusion

# References