Single Molecule Neural Net (SiMoNN)

Joshua Yoon, Maurice Lee

Abstract: Here we discuss the use of deep learning to extract as much information from single molecules detected on a camera. In SMLM, we rely on individual molecules to emit fluorescence in a stochastic manner in order to provide an accurate estimate of its position. By using a simple, fully connected neural network, we were able to properly classify images as "good" or "bad" based solely on their intensity profile with high accuracy. To further explore the possibility of determining the spatial 2D location of each molecule, we trained a YOLO network using a relatively small number of grayscale images, but had very limited success.

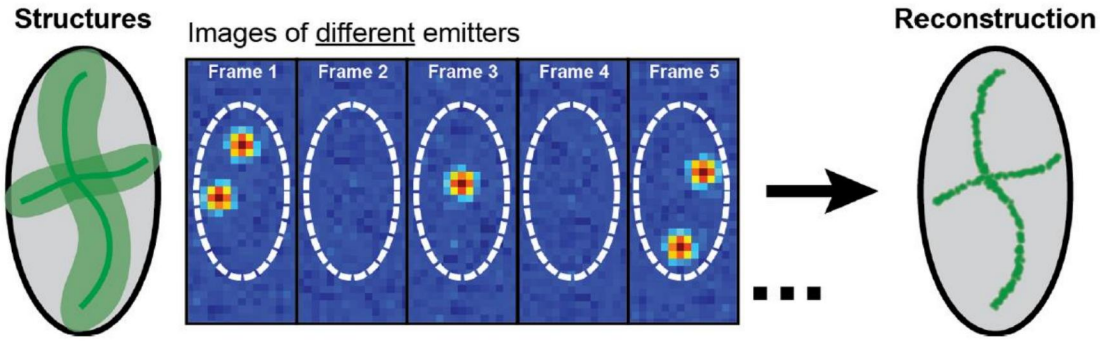

Figure 1: Idea behind SMLM, where we are able to create reconstruction of the structure of interest by causing only a few molecules to be on at a time.

Introduction

Our ability to observe small, nanometer-sized objects is largely limited by the diffractive nature of light. However, if we excite only a few molecules at a time using a powerful laser, we are then able to generate a pointillistic map of the structure of interest (Fig. 1) [1]. For over a decade, this technique known as Single Molecule Localization Microscopy (SMLM) has produced images that have greatly enhanced the resolution to about 10-20 nanometers, about an order of magnitude better than conventional microscopy techniques [3]. However, current methods still suffer from two main issues: (1) slow results, due to heavy computational steps when analyzing data, and (2) throwing out a good portion of our recorded data that is mostly due to overlapping molecules and/or poor signal to background ratios. These two limitations are what motivated us to incorporate a neural network scheme in order to create a simple, fast, and accurate model that produces comparable results with current methods. We created 3 models in our study: (1) ANN, (2) CNN, and (3) a YOLO network. Overall, our results demonstrate reasonable success when classifying individual emitters, but if we try to implement a localization scheme, things get a bit tricky.

Data

Our data consisted of blinking molecules that reside within the nuclei of mammalian cells. We then cropped out a 7x7 window of each individual emitter and performed a 2D Gaussian fit in

order to determine the following features: amplitude, xy position, Gaussian widths, and background offset.

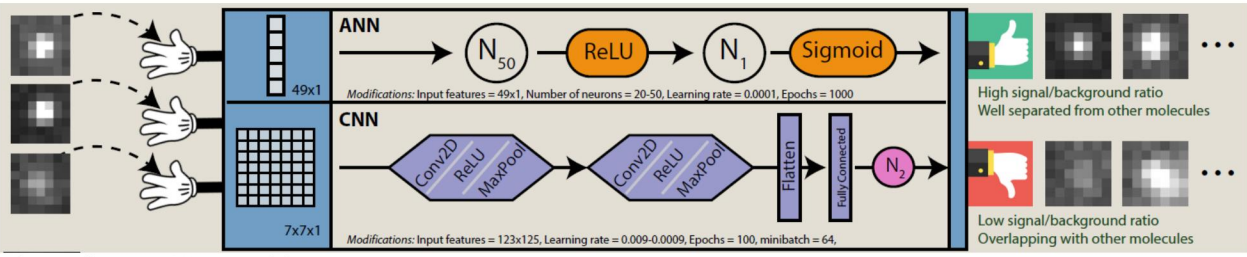

Figure 2: ANN/CNN architectures

ANN: We first classified molecules as "good" or "bad" based on their Gaussian width ($0.8 < \sigma < 1.05$). A "good" molecule is considered to have both high signal/background ratios and is well separated from other nearby molecules. "Bad" molecules are considered to have the opposite characteristics. We then resized our 7x7 images into 49x1 vectors and fed forward ~10000 different images through SiMoNN with a layer of 50 neurons and a final layer of one neuron, where a sigmoid activation function determined whether or not the image was good [4]. We used a logistic loss function, which was used to update all of our weights and biases. Overall, it had high success rates, with a very low test error of 0.01%. Our network does a very good job of detecting good" images, but not as well for the "bad" images.

CNN: In order to see if we could take a different approach towards the same goal, we fed forward our 7x7x1 images through two Conv2D-ReLU-Maxpool modules, flattened the output, went through a fully connected layer, and a final neuron output determined which class the image belonged in: "good" or "bad" [4]. Overall, the network does a better job of categorizing the training set, but does worse with the test set. We used a learning rate ~0.009, ran 100 epochs, with a minibatch size of 64.

YOLO: To make an attempt towards creating a network that could perform localizations, we created a YOLO network where we fed in 123x125x1 images with 7x7x1 annotations that are boxed around the localizations of each emitter [2]. With >5600 epochs, minibatch size of 4-32, momentum = 0.9, and decay rate of 0.0005, we were able to train a model where the loss was ~5, but the bounding boxes in our output were largely rubbish, although a few of the localizations were accurate.

Model	Train Error	Test Error	Tr/Te Split %	Total # Img
ANN	0.34%	0.01%	75% / 25%	10000
CNN	0.20%	0.68%	75% / 25%	20000
YOLO	Very high*	Very high*	99% / 1%	998

Table 1: table of error calculations

Conclusion

The feed-forward neural networks, ANN/CNN, were successful at categorizing different molecules, but if we want to produce a model that produces localizations with sub-pixel accuracy, we'll need to (1) train more or (2) modify the network architecture. YOLO also takes a long time to train, especially if it's done from scratch. Overall, we learned some valuable lessons and have a better grasp of deep learning than ever before. We hope that we can apply these tools to tackle problems such as overlapping emitters and 3D localization of exotic emission profiles.

Bibliography

[1] Betzig, E., et.al. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313, 1642-1645 (2006)
[2] Redmon, J., et.al. You Only Look Once: Unified, Real-Time Object Detection. *Arxiv* (2015)
[3] Diezmann, A., Shectman, Y., Moerner, W. E., Three dimensional localization single molecules for super-resolution imaging and single-particle tracking. *Chemical Reviews*. (2017)
[4] Ng., A., et.al. *CS 230 Notes & Coursera* (2018)