1 Project title

Deep Learning to select particles in electron microscopy images

2 Brief Description of Design Project Goals

Overview: Graduate students and I have worked on 3-D image reconstruction problems for single-particle cryo electron microscopy, e.g., Zheng, Wang, Doerschuk, “Three-dimensional reconstruction of the statistics of heterogeneous objects from a collection of one projection image of each object”, Journal of the Optical Society of America Series A 29(5):959-970, May 2012¹ and Ma, Gong, Aubert, Turker, Kao, Doerschuk, Wiesner, “Surface-tant micelle self-assembly directed highly symmetric ultrasmall inorganic cages”, Nature 558:577–580, 28 June 2018². The applications are both to biological particles, especially viruses, and to nanotechnology particles. The electron microscope provides large images, e.g., 4000 × 4000 pixels, which show many instances of the particle of interest. The first stage of the computational pipeline is to extract small subimages, e.g., 200 × 200 pixels, that contain individual particles. This problem is called “particle picking”. This task is complicated by the low signal-to-noise ratio of the images and the complicated “non-particle” content of the images. The task is simplified by the fact that the different instances of a particle are of the same size.

The goal of the project is to use modern Machine Learning techniques, especially Deep Learning techniques, to solve the particle-picking problem. Conceptually, there are three qualitative steps that are necessary: recognize each object in an image, classify each object that is recognized, and draw a bounding box around each object. There was a workshop³ in April 2018 that considered Deep Learning for a variety of problems in biological single-particle cryo electron microscopy, including particle picking. There is also a github page on all sorts of cryo electron microscopy software, including particle picking software⁴. Please note that the list of particle picking software on the github page is probably incomplete because many of the large packages (e.g., Relion, EMAN2) include particle picking. I think we can improve on what was described at the workshop and what is listed on the github page!

Specific MEng Contribution: Use widely-employed Deep Learning software systems with some additional software “glue” (probably written in Python) to create a software system that achieves particle picking. I would like to write a paper for a computer vision conference like CVPR!

¹http://dx.doi.org/10.1364/JOSAA.29.000959
²https://doi.org/10.1038/s41586-018-0221-0
³http://nramm.nysbc.org/4-10-2018-a-workshop-on-deep-learning-for-cryoem/
⁴https://github.com/barrykui/awesome-cryoem
ECE Field Advisor Name: Peter C. Doerschuk

- Email: pd83@cornell.edu.
- Phone: 607-255-4179.
- Office: Phillips 305.

Outside Field Advisor Name (if applicable):

- Email
- Phone
- Office

Project Web Site:

Number of MEng Students Needed: A team of perhaps two.

Required Skills: The project needs students with a balance of “algorithms” interests, especially Deep Learning and statistical, and “computation” interests. Hopefully, the large majority of the software will be written using high-level packages, like PyTorch or TensorFlow, in Python. Prior experience with research, not necessarily using these tools, is good!

Estimated Project Time Frame: 2018-2019 Academic Year, Two (2) Semesters, 6 credits.