# CS 491 Undergraduate Research

Sergio Servantez
Computer Science Department
Illinois Institute of Technology
Chicago, Illinois 60616
sservantez@hawk.iit.edu
Advisor: Dr. Zhiling Lan

**Report 1: 2/3/19**

This project is based on earlier work [1] investigating deep neural network (DNN) training on large scale computers. This earlier work demonstrated that the training times of these neural networks could be significantly reduced by massively parallel processing. We aim to reproduce these experiments and investigate the same implementation on the Theta supercomputer at Argonne National Lab.

Much of the first two weeks has been spent becoming familiar with the work cited above. This process began by reading the paper and understanding how batch size and communication cost effect performance. Since internode communication is a significant factor in training times, understanding Theta's architecture, especially the Cray Aries network, was a crucial part to the preliminary research. The next major step is to install the parallel DNN implementation on Theta. The implementation was released in the Intel distribution of Caffe (version 1.0.7). First, I obtained the source code so I could become more familiar with the project bundle. The current release of Caffe is version 1.1.4. I have been investigating what changes, if any, have been made to the parallel implementation since the 1.0.7 version. We must first determine which of these version would be best to install on Theta. I have never previously worked with Caffe so I am currently becoming familiar with the framework using tutorials provided by Intel.

Over the next two weeks I have two main goals. First, I would like to become comfortable working with the code from the Caffe framework and gain a better understanding of the higher level structure of the project bundle. The contents of the framework are rather large. Second, I would like to begin the Caffe installation process on Theta. I am expecting this process to be quite difficult.

[1] Yang You, Zhao Zhang, Cho-Jui Hsieh, James Demmel, and Kurt Keutzer. 2018. ImageNet Training in Minutes. In ICPP 2018: 47th International Conference on Parallel Processing