**Diagnosing deep-learner digital surface model evaluators using an evolutionary algorithm on Summit**

Mark Coletti and Alex Fafard

Oak Ridge National Laboratory

colettimaj@ornl.gov

**Abstract**

3D surface models can be derived from overlapping 2D images of a scene. This approach can be used to create digital surface model (DSM) of parts of the Earth from over-lapping satellite images. However, this approach is prone to error from clouds, haze, water, shadows, and other obstructive artifacts. Humans can be used to detect these errors such that other, hopefully error-free, image pair candidates can be tried; but this manual approach can take several weeks given the tens of thousands of DSMs to be evaluated. This process can be automated through the use of Convolutional Neural Networks (CNN), thus shortening evaluation time to a matter of days instead of weeks. An evolutionary algorithm (EA) can be used to improve CNN design accuracy.

An EA can also be used to diagnose CNN problems, such as damaged training data. Several runs on Summit where an EA evaluated thousands of DL configurations yielded no overall improvement in DL performance, which suggested problems with the DL or the corresponding training and validation data. We hypothesized that Contrast Limited Adaptive Histogram Equalization (CLAHE) enhancement that was applied to previously generated DSMs had damaged the training data. Subsequent runs without CLAHE enhanced data yielded significantly improved DL performance.

However, though the DL validation accuracies had greatly improved, they still exhibited consistent sub-par performance, which suggests further problems with the training data and DL approach need to be uncovered and corrected.