Tutorial 3 – Training the deep learning model using the VACC

Now that you have prepared your training data, it is time to train the model using the Vermont Advanced Computing Core, or the VACC. The VACC contains a neural network that allows for efficient training for machine learning. Using our imagery and labels as inputs, it will produce a .h5 file that will be used to classify satellite imagery.

This tutorial picks up from Tutorial 2. At this point, you should already have access to the VACC, and have uploaded your training data.

1. Connect to the VACC using `$ ssh uvmnetid@vacc-user1.uvm.edu`, where uvmnetid is your UVM NetID.
2. `$ conda activate impervious`

Running this command shifts your terminal environment to "impervious" allowing you to start training your model to detect impervious surfaces. We will now run a command that splits our raster into smaller tiles of data that will train the model.

3. `$ python pipeline.py data`, where data is the folder you set up in Tutorial 2 containing your imagery and labels.

We will now take this processed data and add it to the queue in the VACC. The VACC is a resource shared by the UVM community, so it may take days for your job to make it to the top of the queue and execute.

4. `$ ssh dg-user1`
5. `$ cd src/bash_scripts/`
6. `sbatch unetMSE_sbatch.sh`

To check the status of your job, use `qstat`. When the job is complete, the model will be saved to the models subdirectory.