**Readme for High-Low Pressure Image Segmentation:**

**Christina Kumler-Bonfanti**

The high-low pressure UNet models take GFS pressure data and label the images with bounded boxes where either high- or low-pressure systems are located. The labels are created from an algorithm that identifies the pressure system from hand-drawn analysis on GFS data by the FIQAS group. The code will point to where the data (used to be) is stored and assign numerical categories for either low, 1, or high, 2 to produce segmented images for the UNet. The “bounding box” parameters can be specified in this code and the center will be the labeled center of the category. The UNets can be adjusted in structure and other parameters. They output

All code is located on Hera:

/scratch1/RDARCH/rda-ghpcs/Christina.E.Kumler/machine\_learning/ddrf\_low\_tracker/cyclone

Data stored in was stored in "/scratch2/BMC/public/retro/fiqas/aws/%Y/%m/%d" but it has been moved since the conclusion of the initial project.

FIQAS data are in grib2 format.

Code to processes the data:

./prepare\_data\_ddrf.py

This has routines defined in:

../data/fiqas/read\_gfs.py

./label\_data\_ddrf.py

To call this on HPC with specified date ranges for the testing, training, and validation datasets, use the following script:

run\_fiqas\_prep.sh

Different versions of processed files for the high-low labeler are stored in the FIQAS directories "../data\_ddrf/\*" where the naming convention is fiqas-block\_[train/test/val]\_YYYY.zip

Keras scripts to train models based on bounding boxes sizes, which are determined by number of pixels in a model gridcell:

keras\_segmentation\_fiqas\_multi\_cats\_bb.py

This has UNet architecture routines defined in:

../learning/unet.py

../learning/losses/losses.py

Models are saved in the “./models/models\_keeprs/\*json” directory as json files and the corresponding weights are located in the “./weights/weights\_keepers/\*h5” as hierarchical data files.

Plotting/testing validation code on a trained model:

./test\_fiqas\_model\_multi\_cats.py

Outputs are images and corresponding npy files. They are found in “./output\_ddrf\_multi\_cats\_\*”

For running on HPC, the csh will open the singularity environment and then call the sh scripts.