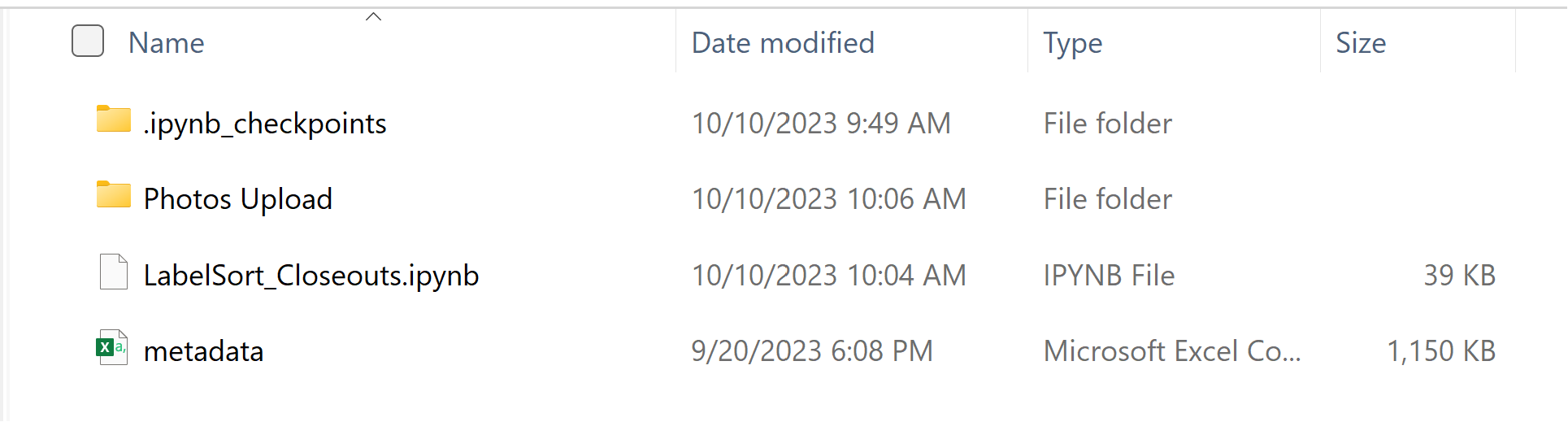
**Roadrunner – setup your environment to match the UChicago student team.**

Installation requirements:  
Jupyter Notebook + the six Python packages included in the requirements.txt package (pandas, os, shutil, numpy, cv2 and PIL; numpy, cv2 and PIL only needed for alligator, blurry and dark models respectively)

Directory setup

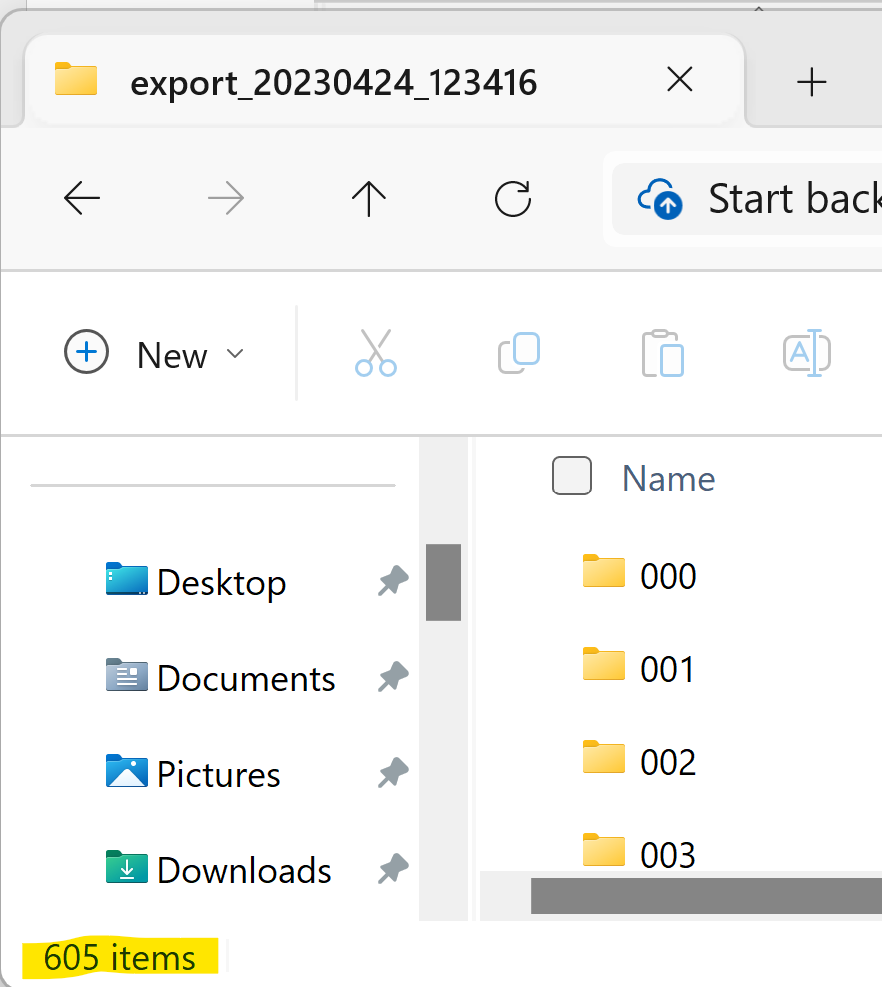
For any of the Azure models, match your local directory to the setup from the shared Google Drive. Here is an example with the closeout image model:



Place a copy of the export\_20230424\_123416 folder inside the Photos Upload folder

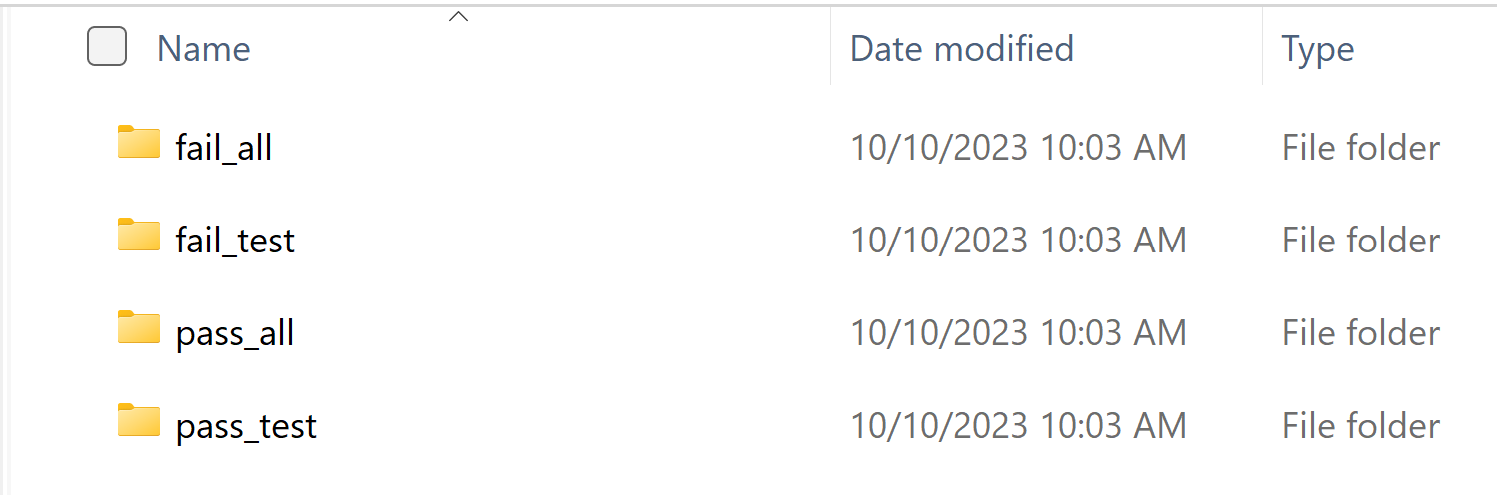


The export\_20230424\_123416 folder should contain 605 numbered sub-folders with images files inside of them.

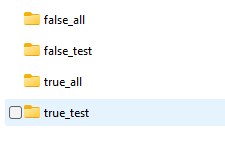


Code running

Once this is setup, open the .ipynb file and run all of the code. This will re-sort the images into folders corresponding to the assigned label for the desired model. In the case of the closeout image model, once the code is run we will have replaced the numbered sub-folders with folders for pass\_all, past\_test, fail\_all, and fail\_test image samples. The pass\_all and fail\_all folders represent the model training sample, which can be uploaded to the Custom Vision platform and tagged accordingly (pass or fail) as training data. The \_test folders are subsequent testing samples used to verify out-of-sample accuracy.



Note that the subfolders are named differently for the alligator model. Instead of fail, it is false and instead of pass, it is true. Images that are alligator freight are in the true\_ folders while those that are not alligator freight are in the false\_ folders.



All of our data samples include more pass images than fails, but we found the Custom Vision models performed best with balanced datasets, so we down-sampled the number of pass images for each. Here are the photos used to train the final versions of each model

**Closeout Pass/Fail Model:**

-all available fail\_all images (542 images)

-000-0 through 556-0\_1 from pass\_all (542 images)

**Non-Closeout Pass/Fail Model:**

-all available fail\_all images (1,042 images)

-000-0 through 788-0 from pass\_all (1,042 images)

**Blurry Model:**

-all available fail\_all images (215 images) plus 135 artificially blurred images (827-0\_11 through 844-0\_3 from pass\_all) for 350 total images. Note that these images are never used in training as pass images.

-000-0 through 064-0 from pass\_all (350 images)

**Dark Model:**

-all available fail\_all images (146 images) plus 200 artificially darkened images (817-0\_1 through 843-0\_4 from pass all) for 346 total images. Note that these images are never used in training as pass images.

-000-0 through 061-0 from pass\_all (346 images)

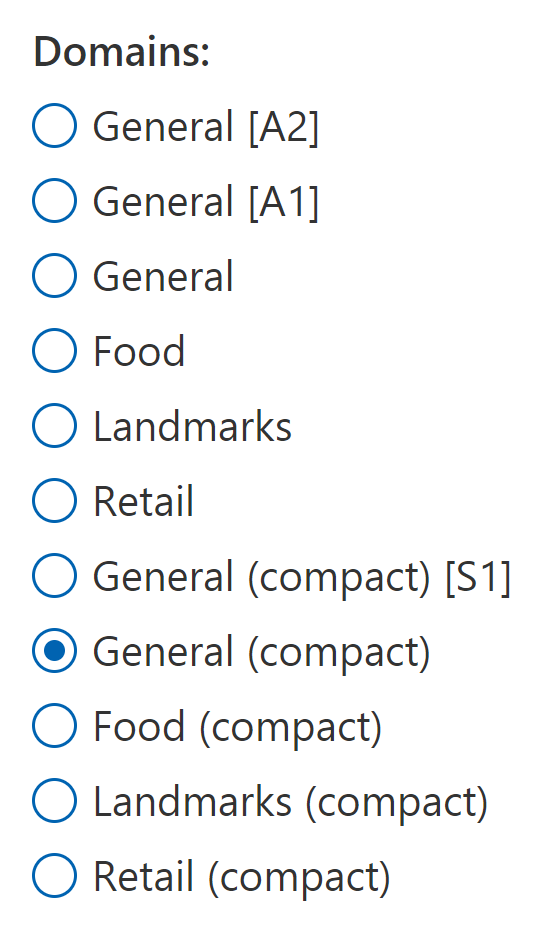
**Alligator Model:**

-all available true\_all images (390 images)

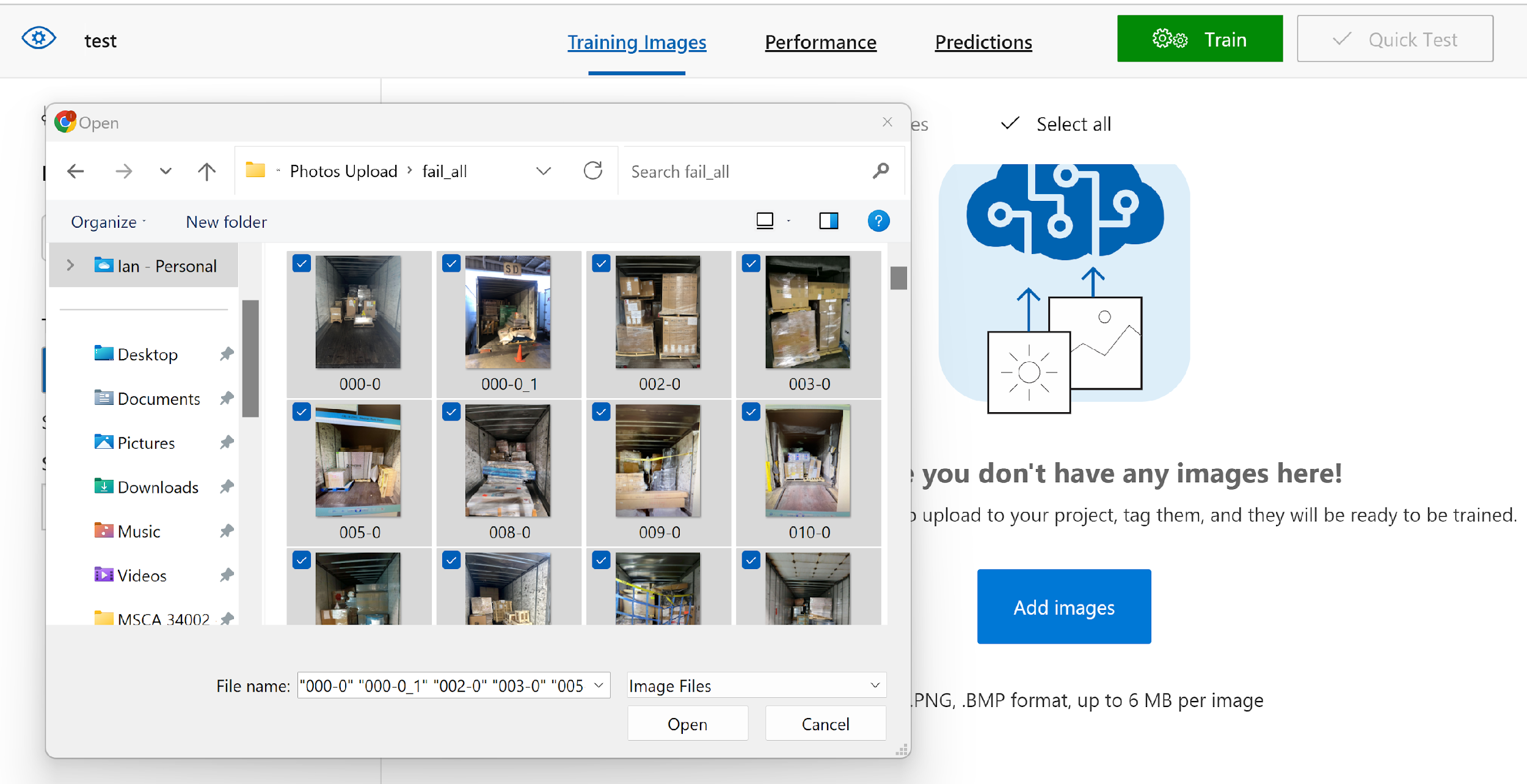
-000-0 through 124-0\_1 from false\_all (390 images)

Azure training

We utilize the customvision.ai portal for model training. A new project is created and all settings are left on the default except the domain is changed to General (compact).



Images can be uploaded by folder through the Training Images tab Add Images button and tagged accordingly.



Once training samples are fully uploaded, models were trained with 24 hours of budget and then fed testing images to assess out-of-sample accuracy. Results were included with the Project Summation document we sent previously.