PROJECT TITLE: Hummingbird Classification

PROJECT SUMMARY (Brief description of the goal, key analysis plan, key data feature- especially if the data is complex)

The purpose of this project is to build a hummingbird species classification model. During this project I will collect hummingbird images, create image pre-processing code for the data set, create a binary (male/female) bird classifier, create a species classifier, and use the classifier to predict future bird images.

MILESTONES (Bullet point STEPs in your project. You can tag “DONE” for things are done)

Project Timeline:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Task Week | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Initial Project Proposal | | |  |  |  |  |  |  |  |  |
| Collect/classify photos from the web and my camera | | |  |  |  |  |  |  |  |  |
| Build data cleaning / hummingbird image processing tool | | |  |  |  |  |  |  |  |  |
| Create binary male/female classification model | | |  |  |  |  |  |  |  |  |
| Create multi-categorical species classification model | | |  |  |  |  |  |  |  |  |
| Draft Final paper / GitHub repository / Kaggle uploads | | |  |  |  |  |  |  |  |  |
|  | | | | | | | | | | |
|  |  | Done/In-progress |  | Overdue | | |  | Planned | | |

PROPOSED ‘TO DO’ FROM THE LAST WEEK (Copy & Paste from your previous week’s TO DO)

1. Create a bird/no bird binary classification model to help sort the images collected. - complete
2. Create bird directory so images can be moved into the correct slot as they are received. - complete

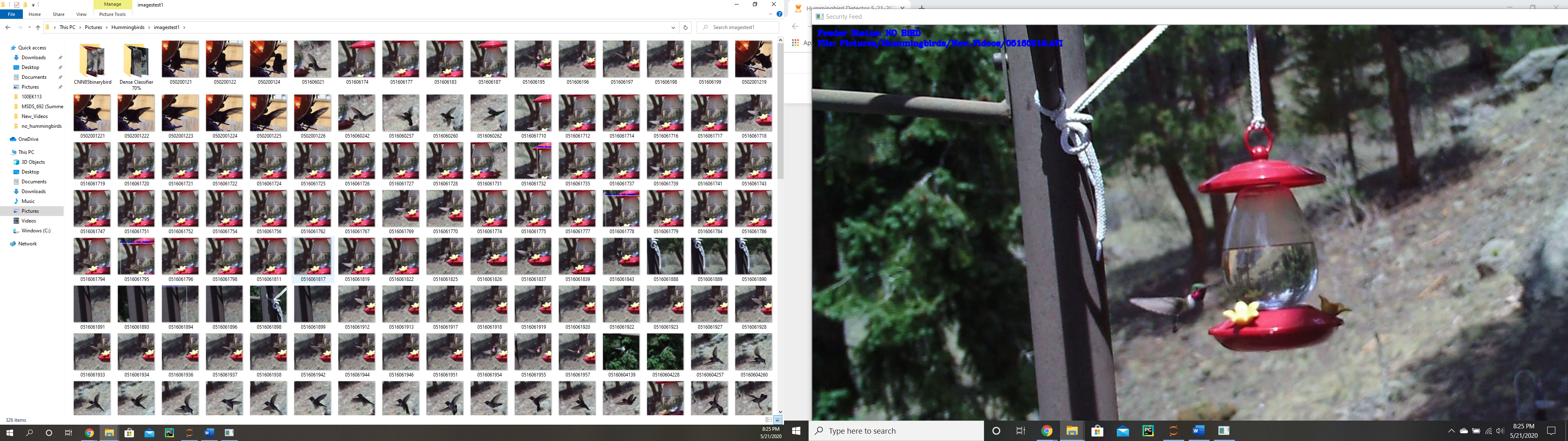
THIS WEEK’S PROGRESS (Give bullet points and briefly explain what you accomplished or dealt with some of the milestones during the week)

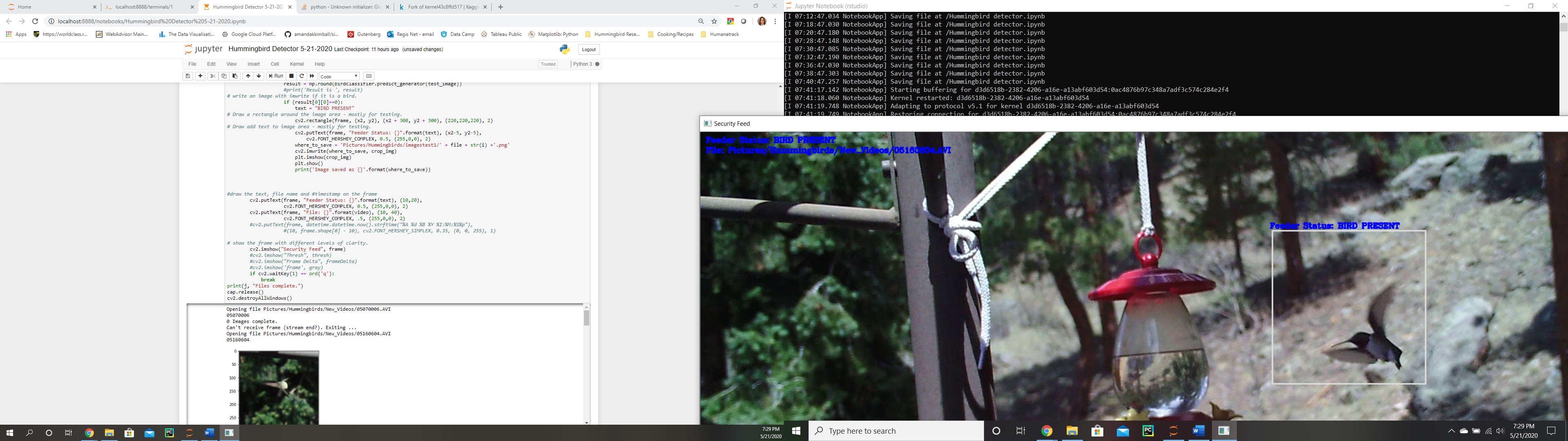
1. I continue to build my collection of hummingbird images. Images are coming into my camera, I built the directory structure for each species and have started populating. This will be my focus in week 4.
2. I added the following information to my github repository:
   1. New bird finder binary classification model – Jupyter notebook.
   2. I also updated the ‘Hummingbird detector.ipynb’ with the bird finder model to sort through the images.
   3. I was unable to upload my birdclassifier.h5 file because it is 810MB.
   4. Here is the link: <https://github.com/amandakkimball/Hummingbird-Data-Collection-Classification>

ISSUES AND DISCUSSION (Bring up any difficulties and things to discuss. Also, send me a reminder if you want to discuss sooner)

This week I focused on a binary classification model to identify a bird in my video feed. (10 second clips for now but could be a continuous feed once everything is setup correctly.) The classification was a little frustrating at first. The videos from my camera are not the high resolution pictures I’ve seen in some of the image competition projects on Kaggle. I used an existing dataset the 100-bird-species [2] (now 200) with the Annas and Ruby Throat Hummingbird directories and the 40 images that I collected from my camera last week as my bird present images. I also used the not humming bird images from last week and tried to increase this dataset with images of flowers from the web. This turned out to be difficult for training and I ended up going back to just the non-bird images from my camera. I trained a simple dense model and a CNN model for the analysis and after some pretty extensive augmentation was able to get about 80-85% with the CNN and 85% with the dense model. See my Github for details [1]. This has decreased the time to review the video clips that I have created. I have roughly 2000 videos to go through.

Here are some screenshots from my analysis: It’s exciting to see many bird images with a few non-bird images from my video processing process.





TO DO (Give bullet points and briefly explain your plans for the next week)

1. During week 4, I will focus on populating my images in the directories for each bird species/gender. With the tools created it should go much more quickly but I have compiled several 1000 – 10 second video clips and a few 100 bird images that need to be pre-processed/cleaned.
2. As time allows, I will try to optimize my binary bird detector model now that I have more birds from my camera and with a consistent background.

RESOURCE (Optional: list resource or links you want to share with me)

[1] <https://github.com/amandakkimball/Hummingbird-Data-Collection-Classification>

[2] https://www.kaggle.com/gpiosenka/100-bird-species