Purple Martin Video Processing

FINAL REPORT

Under the guidance of

Dr. Robert DeLeon, PhD

Email: rldeleon@buffalo.edu

by:

Yashas J Shamaraju

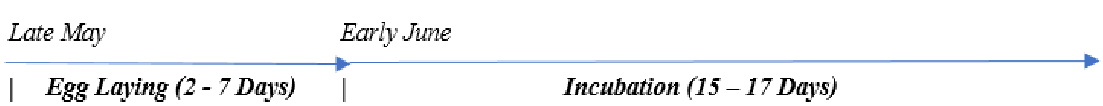
Introduction

The purple martin (*Progne subis*) is the largest North American swallow. They are known for their speed and agility in flight, and when approaching their housing, will dive from the sky at great speeds with their wings tucked. The average length from bill to tail is 20 cm (7.9 in). Adults have a slightly forked tail. Adult males are entirely black with glossy steel blue sheen. Adult females are dark on top with some steel blue sheen and lighter underparts. Subadult females look similar to adult females minus the steel blue sheen and browner on the back. Subadult males look very much like females, but solid black feathers emerge on their chest in a blotchy, random pattern as they molt to their adult plumage.

Figure 1: Female Purple Martin

Figure 2: Male Purple Martin

After spending a few months in Brazil, Purple Martins will begin their long migration back to North America. Copulation starts in the beginning of May. Egg laying commences after copulation occurs. Females lay one egg per day, usually in the morning, for a total of two to seven pure white eggs. After the penultimate, or next to last, egg is laid, females begin incubation. Only females can incubate eggs because only they have a brood patch, a featherless area rich in blood



vessels that transfers heat to the eggs. Males may insulate the eggs for short periods of time while the female leaves the nest. Fifteen to sixteen days later, the eggs begin to hatch. The eggs may not hatch on the same day, but rather it can be spread out over two or even three days. Once the young have hatched, both parents begin feeding the young. Pin feathers and downy feathers begin to emerge after 7-10 days.

Conservation

The New York State Purple Martin Project is a collaborative effort of NYS Ornithological Association educating citizens in the conservation of Purple Martins, with the goal of increasing the declining Purple Martin populations throughout New York State. In New York state, purple martins have declined by 39% since 1985. Martins can carry a number of parasites that can cause bird mortality (especially in the freshly hatched). The parasites often aren't found on the birds but at the bottom of their nesting material. In order to study the effects of parasites on the Purple Martins mortality, we have to monitor the nests of these birds.

Therefore, cameras were installed in multiple Purple Martins nest to record the activities in the nest. We have 8 channels (or 8 different nests), and the video is recorded from 6 AM to 10 PM every day from late May to Early July (Approximately 60 days). This results in a huge pile of videos per channel, amounting around 600. This makes it difficult to monitor the videos manually. This is where data science and machine learning methodologies comes in handy.

Objectives

- Design a video processing model that reduces the time taken to process each video. We are left with 4 TB of video to be processed.
- Train a neural network to detect the birds in each frame (also to find if it’s a female or male or none)
- Find the nest attentiveness of the birds by processing each video.
- Plot the nest attentiveness against the time to find out important details in the nesting birds life.
- Superimpose multiple nest data and find the similarities between different channels.

Data

Training data:

- 2512 Images with the birds in the nest and 1103 Images without birds in the nest
- Each image has 1980*1080 resolution.
- 30% of these images are taken as validation data

Test data:

- Each video is read frame by frame to make the test data.

- Each video is named as the ChannelName_Timeframe when it was recorded.
  Example: ch02_20180528191746.mp4 ➔ chXX_YYYYMMDDHHMMSS.mp4
- A 33-minute video is made up of 1980* 30 frames. (videos are at 30FPS).
- To process a 33minute video we need to predict 59,400 frames as incubating and non-incubating.
- So total test set is 59400 * (500 videos/channel) * (8 channels).
- In total 237,600,000 images for incubation and provisioning each.

Figure 3: Image with bird

Figure 4: Image without the bird

Methodology

1. *Video preprocessing*.
   - Since we have 237 million images to predict, we pick an image for every 30 frames. This is done because the bird doesn't move much in under a second. Because we pick one in 30 frames, the number of frames is equal to the number of seconds in the video. After this, we are down to 8 million frames.
   - Each frame is of 1980*1080 resolution (~220 KB). We resize the images to 128*128 (~20KB).

1980*1080 128*128

2. Neural Network Architecture.

1. Convolutional Layer:
   - 3 layers of convolution
   - Filter size 3*3.
   - The first convolutional layer, we pass n images of size width*height*number_channels (3 for RGB), then this has the size [n width height number_channels].
   - Finally, we use a RELU as our activation function which simply takes the output of max_pool and applies RELU.
2. One Flattened layer
   - The Output of a convolutional layer is a multi-dimensional Tensor. It is converted into one-dimensional tensor. This is done in the Flattening layer.
3. Two Fully connected layers.
   - We define a function to create a fully connected layer. In fully connected layer, we take all the inputs, do the standard z=wx+b operation on it.

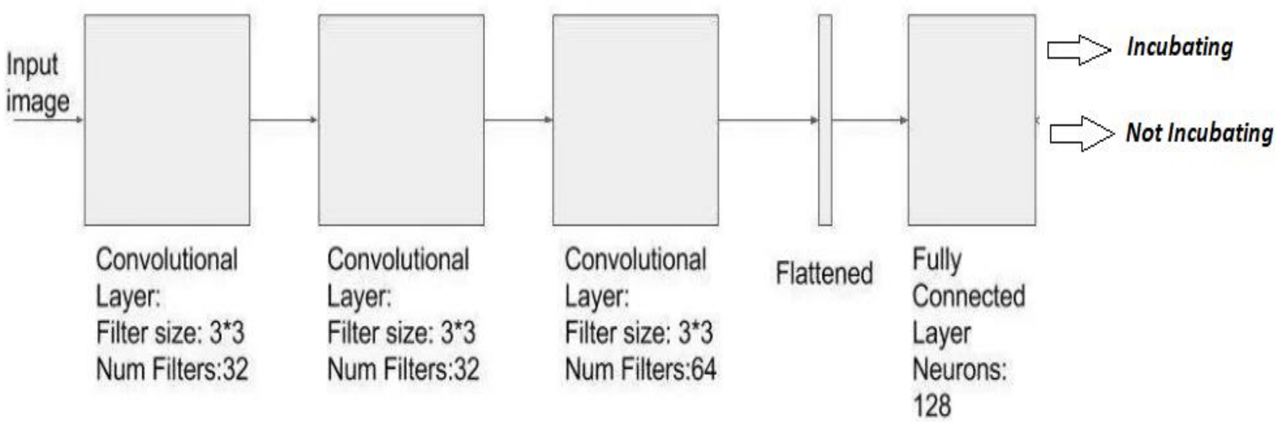


Figure 5: CNN Architecture

3. Optimization

Tensor flow implements most of the optimization functions. We shall use AdamOptimizer for gradient calculation and weight optimization. Learning rate is set at 0.00001.

4. Training

We reach 96.8% accuracy on training data and validation data at around 130 Epoch.

On furthering the epochs, the model starts to over fit.

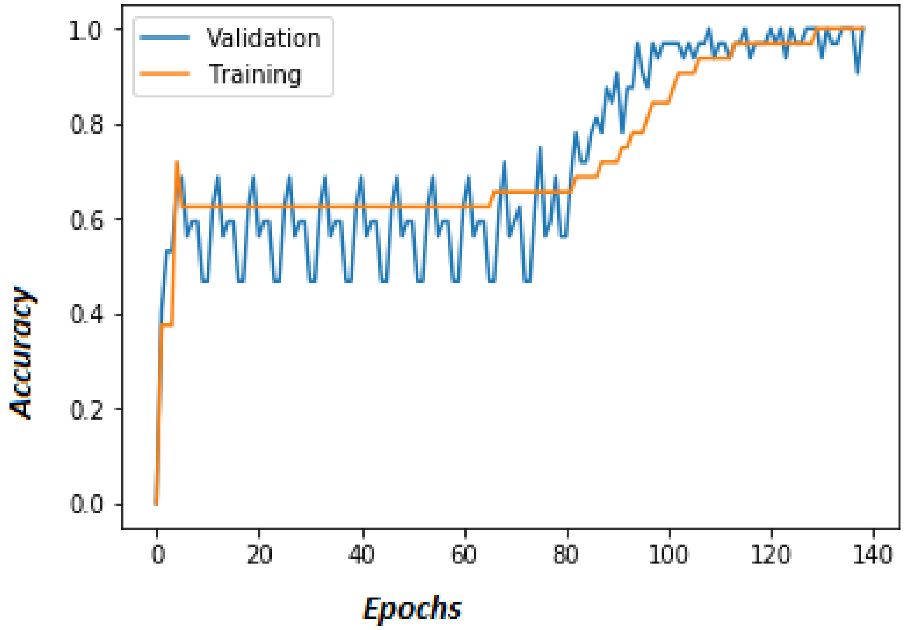


Figure 6: Validation vs Training accuracy

The trained model is saved to be loaded by the predict function for prediction.

5. *Prediction*
    - We get probabilities of 'incubation' and 'not-incubating'. If the probability is greater than 0.5 then it is changed to 1.
    - Example: [[0.99398661 0.00601341]] -> [[1.0 0.0]]
6. *Video listing.*
    - Every video in a folder is sorted and the channel name is stripped.
    - The user has an option to choose the start date and the end date.
    - Based on the time interval given, specific videos will be processed.
    - Nest attentiveness is calculated by this formula.

$$NA = \frac{\boldsymbol{Number\ of\ frames\ with\ birds\ present}}{\boldsymbol{Total\ Number\ of\ frames\ Being\ processed}}$$

- Plot Nest attentiveness vs time. Plot different graphs over each other to find the transition between incubation and provisioning.

Results

Name	Date	Time	Duration	Nest Attenti	prediction
ch09_20180612164703.mp4	6/12/2018	16:47:03	2023	0.4552645	['incubating', 'incubating', 'incubating', 'incubating', 'incubating', 'incubating', 'incubating', 'incubating', 'incu

Figure 6: Output of each video.

NA	Timestamp
0	2018-05-23-06:00:00
0.053386	2018-05-23-06:00:32
0.084486	2018-05-23-06:34:15
0.404844	2018-05-23-07:07:59
0.916502	2018-05-23-07:41:42
0.861097	2018-05-23-08:15:26
1	2018-05-23-08:49:09
0.701928	2018-05-23-09:22:52
0.417202	2018-05-23-09:56:35
0.398418	2018-05-23-10:30:18
0.412259	2018-05-23-11:04:00
0.070652	2018-05-23-11:37:43
0.371725	2018-05-23-12:11:28
0.290657	2018-05-23-12:45:10
0.371231	2018-05-23-13:18:54
0.035591	2018-05-23-13:52:36
0.181414	2018-05-23-14:26:19
0	2018-05-23-15:00:02
0.025704	2018-05-23-15:33:45
0.061789	2018-05-23-16:07:29
0.026693	2018-05-23-16:41:12
0.08498	2018-05-23-17:14:54
0.016807	2018-05-23-17:48:39
0.031142	2018-05-23-18:22:21
0.086505	2018-05-23-18:56:04
0.014829	2018-05-23-19:29:48
0.020761	2018-05-23-20:03:31
0.56044	2018-05-23-20:37:14
0	2018-05-24-06:00:00
0.26087	2018-05-24-06:11:01
0	2018-05-24-06:44:45
0.130005	2018-05-24-07:18:28
0.590213	2018-05-24-07:52:11
0.834404	2018-05-24-08:25:54
0.776569	2018-05-24-08:59:37
0.204647	2018-05-24-09:33:20
0.008403	2018-05-24-10:07:03
0.078102	2018-05-24-10:40:46

Figure 7: Output for channel 9

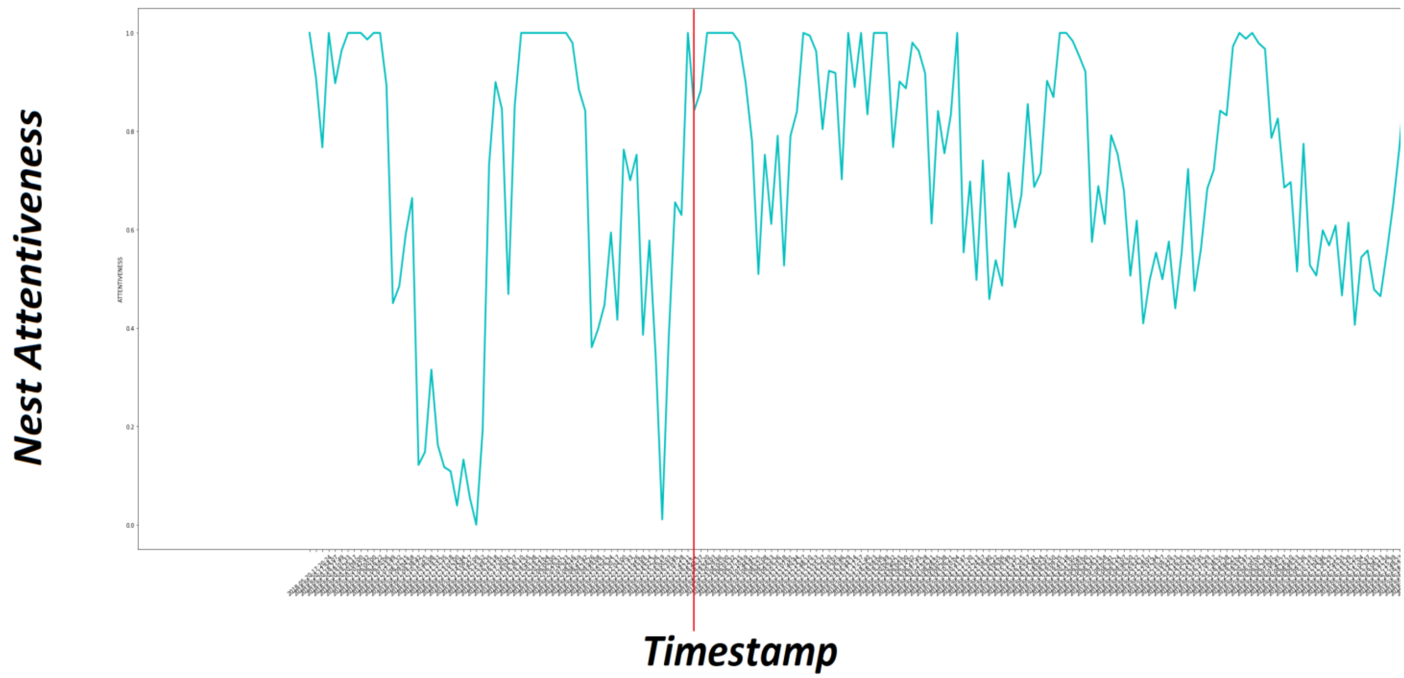


Figure 8: Nest Attentiveness vs TimeStamp()

Inference:

1. We find that videos are recorded from 6 am to 10 pm every day.
2. Birds are active between 7 am to 7 pm with nest attentiveness less than 50%
3. Redline on the graph represents the last day of laying the egg. After which we see a steady increase in the nest attentiveness as incubation starts.

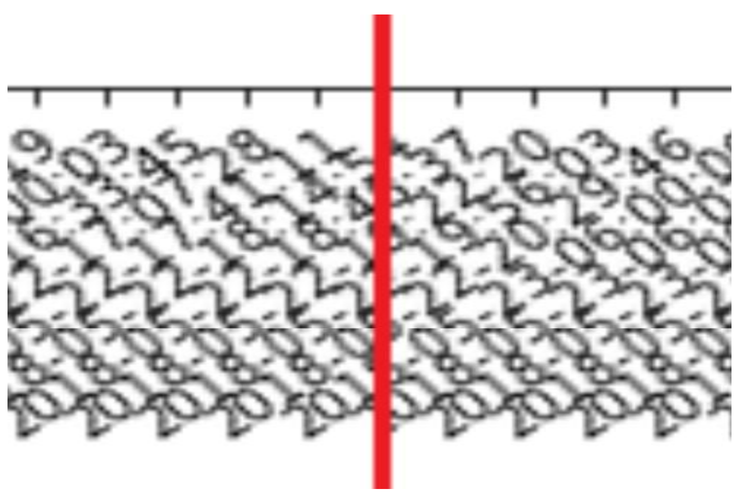


Figure 9: TimeStamp of the last egg laid. (2018 / 05 / 22)

Challenges:

- Poor lighting in the videos caused the model to predict a bird even though there was no bird. This issue was resolved by training the model with poorly lighted images also.

Figure 10: Normal image

Figure 11: Poorly lit image

- There was difficulty in distinguishing between male and female birds and distinguishing between a nestling and a parent. Male birds are can be distinguished by their prominent purple colored plumage, but sometimes even young female birds can have purple plumage. This results have low accuracy in distinguishing male and female birds (~60%).

Figure 12: Male Bird with some glossy steel blue sheen

Figure 13: Female bird with glossy steel blue sheen

Impact

Monitoring an animal is crucial in the process of its conservation. During such monitoring process we may produce a lot of data. Identification by human eye becomes difficult and time consuming. Arriving at a critical mass of data for population analysis can take years (especially for rare or endangered species). Long required observation periods and manual data processing (e.g., matching photos “by eye”) can create multi-year lags between study initialization and scientific results, as well as create conclusions too coarse or slow for effective and optimizable conservation action. This limits the scope, scale, repeatability, continuity, and ROI of the studies as they face the limits of their home-grown tools and IT capabilities.

Wildlife researchers lack a common yet customizable platform for collaboration and often don’t have the technical experience or budget to take advantage of advanced computing tools (e.g., computer vision, artificial intelligence). These projects help in providing an essential tool to mitigate these problems.

Images have become the most abundant, available and cheap source of data. The explosive growth in the use of digital cameras, together with rapid innovations in storage technology and automatic image analysis software, makes this vision possible particularly for large animals with distinctive striped, spotted, wrinkled or notched markings, such as elephants, giraffes, and zebras. This large number of collected images must be analyzed automatically to produce a database that records who the animals are, where they are, and when they were photographed. Combining this with geographic, environmental, behavioral and climate data would enable the determination of what the animals are doing, and why they are doing it.

In future, we may use this model to identify different species using natural markings, genetic identifiers, or vocalizations. Also, we can build an AI system that processes a video sighting to recognize the species and individuals. This can also be used as a knowledge base for people.

Reference:

[1] Wikipedia.org: Purple martin: [Online]. Available from: https://en.wikipedia.org/wiki/Purple_martin [Accessed 8th December 2018].

[2] Wild Me: Gateway to A.I. and Computer Vision: [Online]. Available from: https://www.wildbook.org/doku.php [Accessed 9th December 2018]

[3] NY State Purple Martin Project: [Online]. Available from: http://www.friendsofiroquoisnwr.org/purplemartins/pm_docs/PurpleMartins_012615_C.pdf

[4] Anders Pape Møller: Parasites, Predators and Nest Boxes: Facts and Artefacts in Nest Box Studies of Birds?

https://www.jstor.org/stable/3565628?seq=1#metadata_info_tab_contents

[5] http://cs231n.github.io/convolutional-networks/