Image Classification: Bird Species

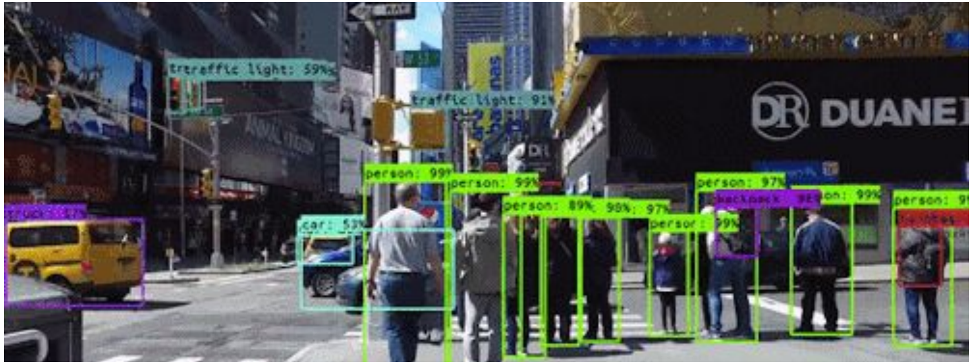
The Principal Components

Shaurik Deshpande, Johanna Niggemann, Mary Hunter Russell, Cam Santos, and Henry Shugart

Motivation

- Why image classification matters:
  - The concept of “**computer vision**” - developing programs that allow a computer to understand what is being presented to it in the form of pixels - is increasing in popularity across tech fields.
  - Image classification is key to developing computer vision.

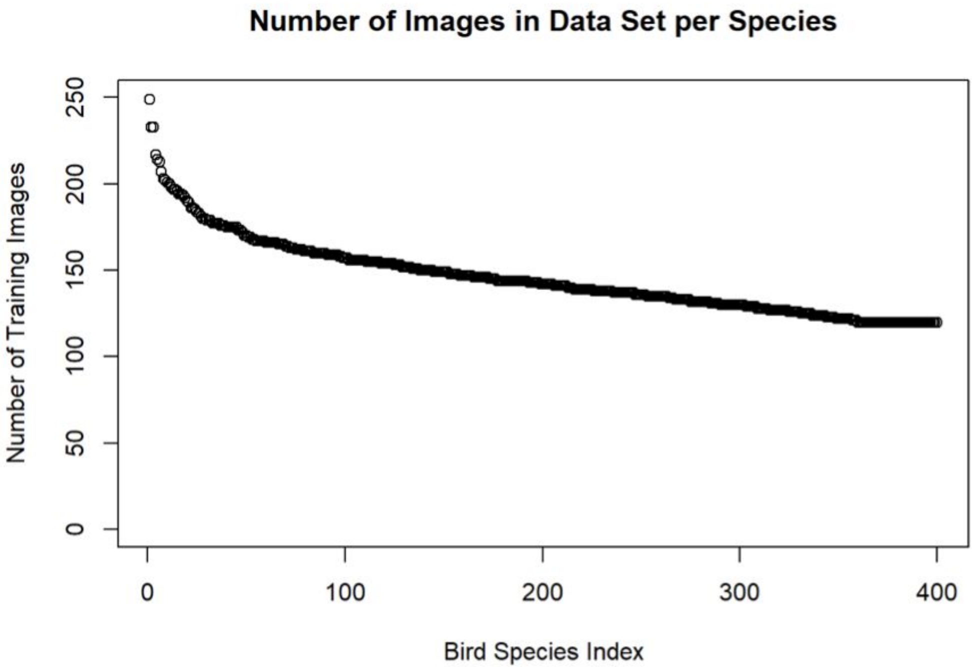
- It has broad applications today ranging from disease detection in medical imaging to programming autonomous vehicles.



Dataset

- The data includes 58,388 observations across 400 species and was already partitioned into training and test data for the model.
- Here we see that there are a variety of poses which each bird takes. This should help improve the model as a picture from any angle should be classified well.

Training Data



- Since the data was pre-split into training and testing segments, we took a look to see how much data we had in each.

- In the training dataset, we have between 100-200 (avg 150) images per species, enough to train classification models.

Resizing Images

Bornean Leafbird

Tailorbird

Flame Bowerbird

224 x 224 x 3

26 x 26 x 3

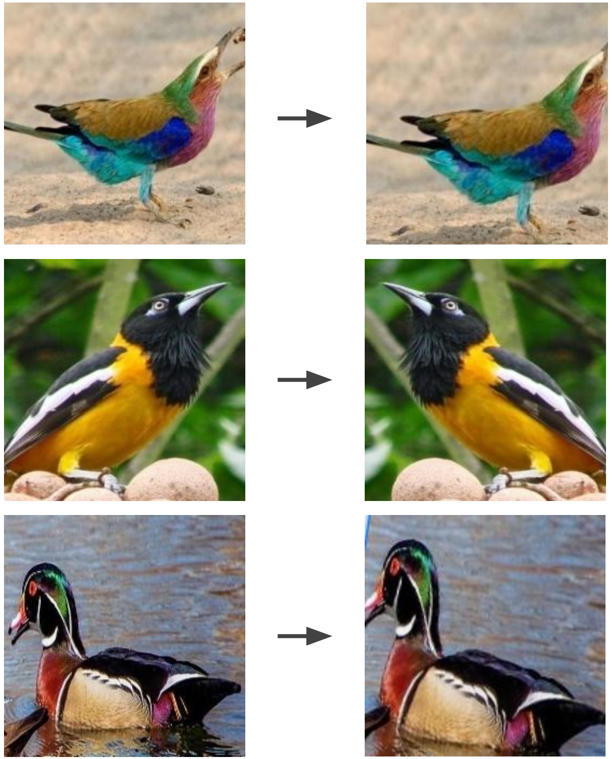
16 x 16 x 3

Data Preprocessing

Zoom

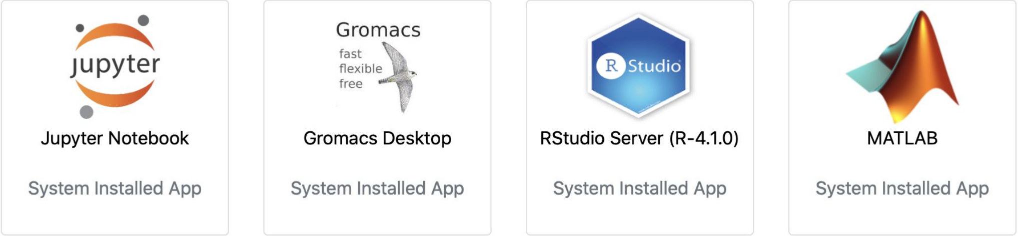
Flip

Sheer



- ❏ Preprocessing images can help prevent overfitting in models
- ❏ We used random combinations of 3 image augmentations on our training data in the hopes of preventing overfitting
  - ❏ Zoom up to 10%
  - ❏ Horizontal Flip
  - ❏ Sheer up to 10%

Longleaf

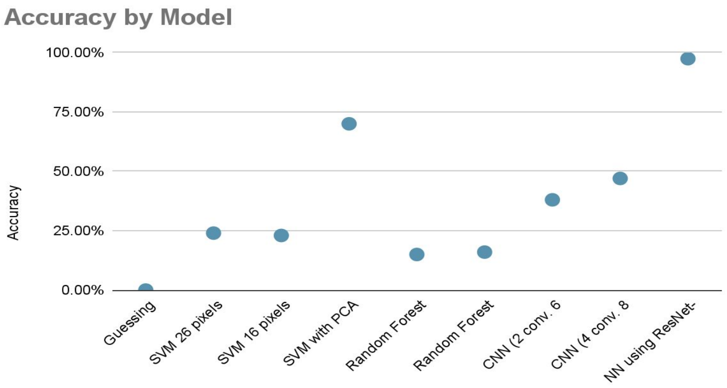


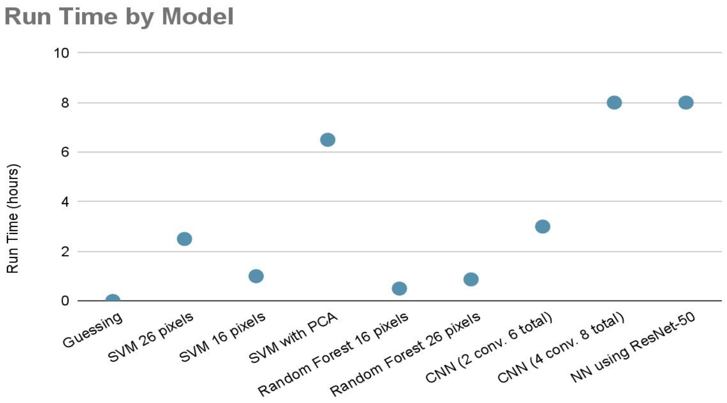
Welcome to OnDemand, a Data Science platform and portal to Longleaf

UNC INFORMATION TECHNOLOGY SERVICES

- ❑ UNC's supercomputing system
- ❑ Targeted for data science and statistical computing workloads, very large memory jobs, and GPU computing
- ❑ Allowed us to use both R and Python to run models in a much more manageable amount of time since our dataset is so large

Summary of Models





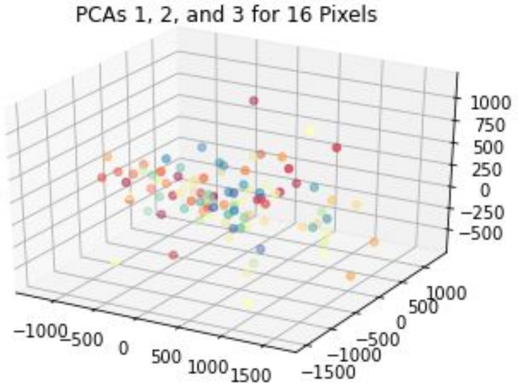
- There is a clear tradeoff between accuracy and run time for our models.
- The most accurate models were the neural networks with many layers, so it is not surprising they take more processing time.
- Even with longleaf, most of our models still took hours to train. However, this is not a debilitating factor as models will still run on new data quickly once trained (without PCA).

SVM

- 16 Pixel SVM Accuracy: 23%
- 26 Pixel SVM Accuracy: 24%
- 16 Pixel SVM with PCA accuracy: 70%*

- Given the values of the SVMs without PCA, it appears that there may not be much separability in the raw data, or that there is a non-linear kernel (which would vastly increase computation time)

- PCA appears to improve the model accuracy a significant amount



Random Forest

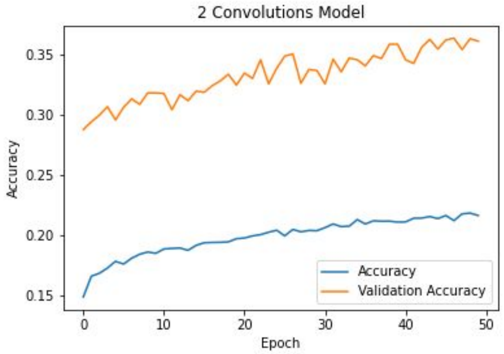
- 16 Pixel Accuracy: 15%
- 26 Pixel Accuracy: 16%

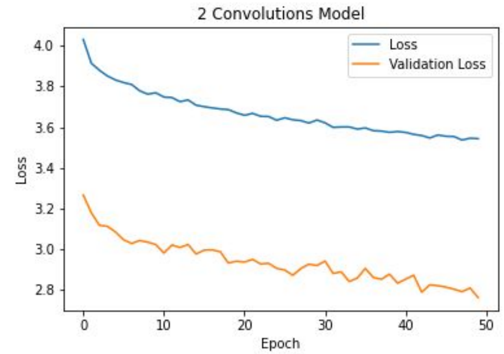
- Although random forests are usually good image classifier, our random forests did not perform well

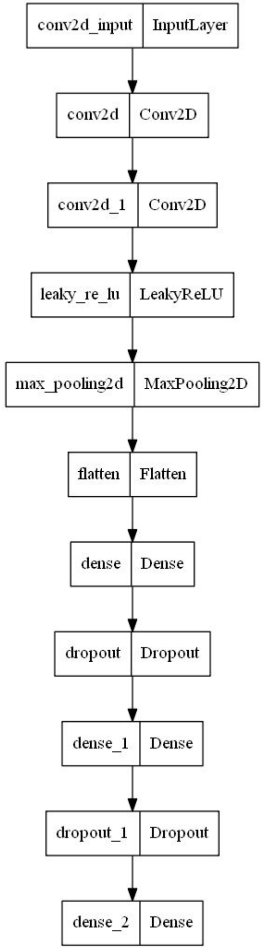
- Performance constraints such as:
  - Important information lost in reduction from 224 x 224 to 26 x 26 pixels
  - Small total number of trees in random forest (100)

NN with 2 Convolutional Layers

- Accuracy: 38%
- Very simple architecture
  - 2 Convolutional Layers
  - Pooling
  - 2 Dense Layers
  - Dropout to prevent Overfitting

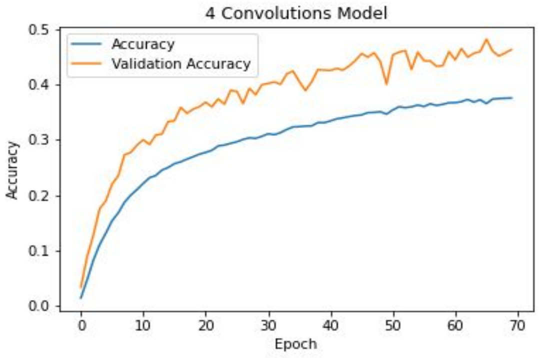


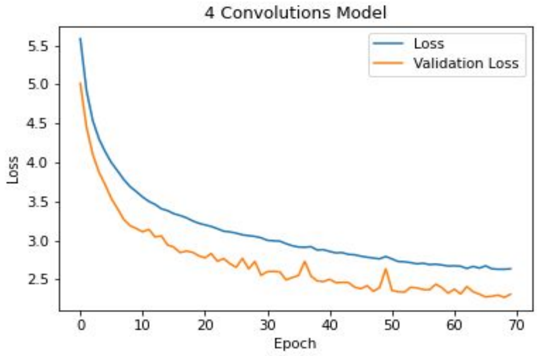


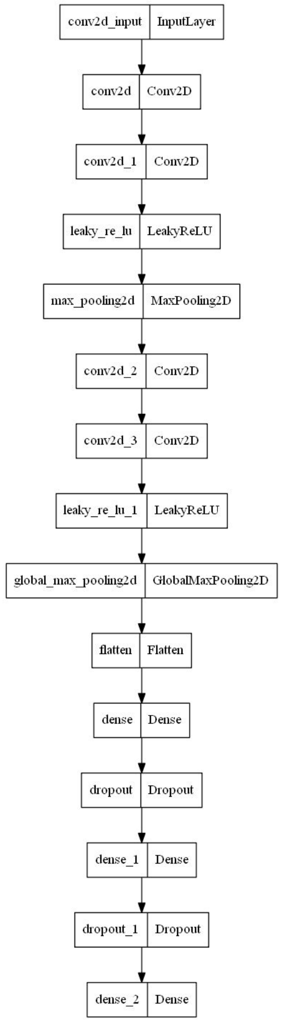


NN with 4 Convolutional Layers

- ❑ Accuracy: 47%
- ❑ Improvement of ~9%, but took more than 2.5x time to train
    - ❑ Added 2 convolutions
    - ❑ Used Larger Images

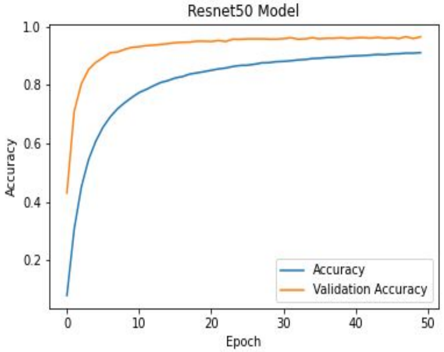


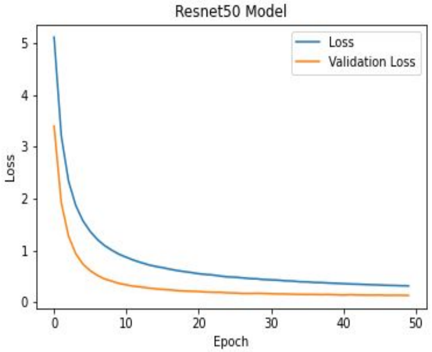


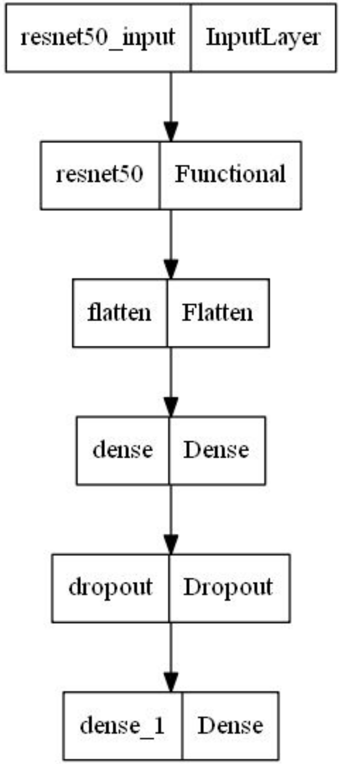


NN ResNet50

- Accuracy: 98%
- Utilizes pre-trained ResNet50 Neural Net
- Uses weights from ImageNet







Takeaways & Next steps

- ❏ Faced challenges in classification since there are so many possible categories (400 species), but any of the models are still an improvement over randomly guessing
- ❏ Our neural network models were the most successful, particularly ResNet which has been pre-trained on millions of images. Tuning this model is likely the best next step
- ❏ Classifying bird species is hard - even humans struggle with detecting the minor differences - so it would be interesting to see how these models work with other datasets (different types of animals, facial expressions, traffic signs, etc.)
- ❏ Next steps: run models with further hyperparameter tuning and additional preprocessing

Q&A

Works Cited

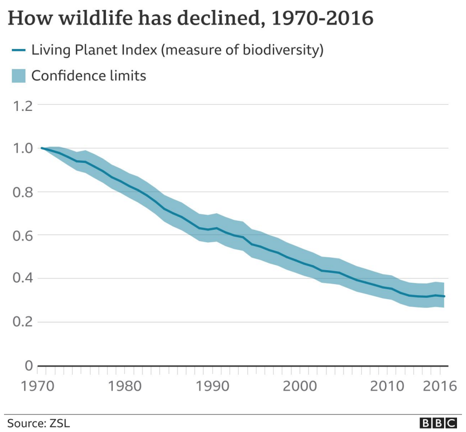
- https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e
- https://keras.io/api/applications/resnet/

Motivation

- Due to habitat destruction and climate change, measuring avian biodiversity is more critical than ever before.

- Traditional monitoring requires intensive human effort to conduct field observations, track movement, and other difficult-to-scale operations.

- Developing a system of classifying species just by images would help ornithologists to

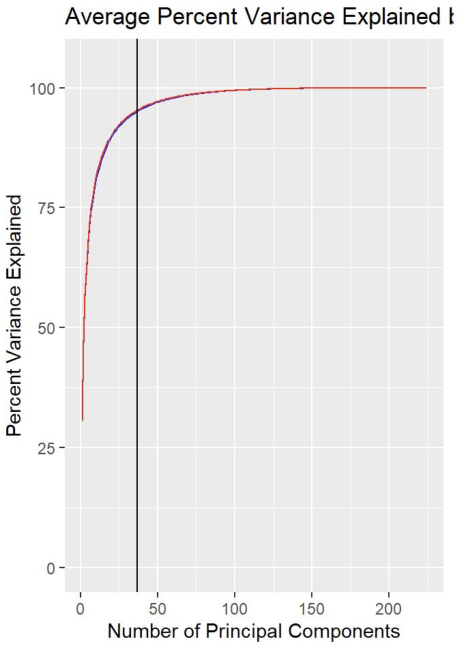


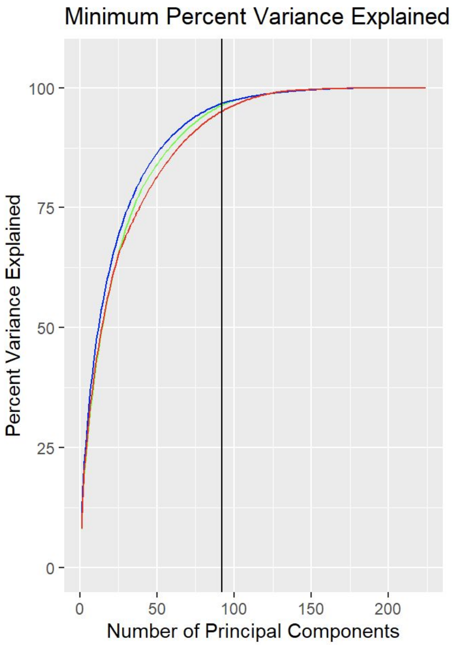
Note about Prediction Accuracy

Warblers of North America Field Guide

- ❑ The baseline prediction by guesswork would be 1/400, or .25%
- ❑ It is difficult for ornithologists to distinguish species without environmental, seasonal, and behavioral context

EDA: Compressibility





- ❑ Since our dataset is so large, we looked to see if we can potentially reduce the size of each image prior to training to make our model training more efficient.
- ❑ We can see that most images are very compressible and even where the data is not easily projected into lower dimensions, there is some ability to dimension reduce.
- ❑ Based on random sample of 300 images

EDA: Image Segmentation

- Next, we looked at a couple of images in the data set and its ability to be segmented using the superpixels algorithm.

- We can see some data is very well segmented (top) while other the algorithm struggles to identify the boundaries of the birds in others (bottom).

- While this was helpful in gaining a preliminary understanding of our dataset, we did not continue to use image segmentation in our project

Summary of Model Accuracies

Model	Accuracy (rounded)	Approx Time
SVM with 26 pixels	24%	2.5hrs
SVM with 16 pixels	23%	1hrs
SVM with PCA	70%	30 minutes + 6hr PCA
Random Forest with 16 pixels	15%	30 minutes
Random Forest with 26 pixels	16%	50 minutes
CNN (2 conv. 6 total)	38%	3 hrs
CNN (4 conv. 8 total)	47%	8 hrs
NN using ResNet-50	97%	8hrs**

Overview of Techniques / Methods

- ❑ Neural networks → multilayer deep learning method to create function(s) for classification
- ❑ Random forests → decision tree model where splits are created with random subset of features from the full model
- ❑ SVM → uses hyperplanes to separate data points into separate classes based on a specific, possibly non-linear, kernel function

Overview

- Motivation
- Data
- Techniques
  - SVM
  - Random Forests
  - Convolutional Neural Network
- Problems
- Evaluation of Techniques