Model Comparison for Bird Species Identification
ISyE 6740 - Spring 2022 Project Proposal

Team Members: Joe Laniado (GTID: 903233267)

April 5, 2022

Contents

1 Problem Statement

Ever since computers were created, we have been trying to find ways for them to automate difficult tasks and make our lives easier. Things like finding the shortest path between two places, drawing a three dimensional floor plan for a new skyscraper, or even doing complex mathematical calculations in seconds; have become as trivial as just opening your phone or laptop and asking a question. One of such questions that came up during the late 1960s was: If we give eyes to a computer, would it know what it was looking at? Of course the answer was no, but then the goal became to find a way to teach it how to do it. This was the birth of the scientific field of computer vision. This discipline deals with analyzing, processing, and uderstanding visual data in a way which allows a computer to see an image or video, and be able to extract information and make conclusions from it. Using X-ray scans for automatic diagnosis of lung diseases in patients, or collecting quantitive information about the traffic flow of a city using only a camera; are some notable applications of this that are currently in use.

After pondering the applications of such computational advancements while doing some bird watching near my home I started to wonder if there could be a way to use computer vision to simplify the process of identifying our feathered friends. I wanted to develop an image classification model that by training it using labeled pictures of different types of birds; if given new input, it could correctly identify which species the new birds belonged to in a quick and efficient way. This process would save me so much time of manual classification of hundreds of pictures of birds in my camera that I just haven't had the time to sit down and check. Fellow local scientists and bird watching enthusiasts have also expressed excitement over the creation of such a model. Furthermore, out in the field spotting a bird could last only a couple of seconds before it flies away, so any time saved on identifiying the species could be used on monitoring other behaviours or collecting scientific information about such a sight. For this a real time video classifier could be used in order to detect and classify the bird before it flies away, although such a project would be a little more advanced than a simple image classifier.

Using the different algorithms learned in class, there is a number of ways to approach the development of our bird species identifier. By means of dimensionality reduction, density estimation, traditional classification algorithms, and deep learning; we will build and compare different image classification models until we find the most accurate and efficient approach to correctly identify the species of each bird. We will evaluate each model with a number of performance variables and analyze their behaviour for this specific classification task. Finally, the best model will be selected and further scalability and functionality implementations will be discussed.

2 Data Source

The original dataset consists of pictures from more than 400 different species of birds. This includes 58388 training images, 2000 test images, and 2000 validation images. When it comes to the images themselves; each picture is 224 x 224 x 3 color images in jpg format, and contain only one bird per image. The bird in the image usually takes more than 50% of the pixels in the image making it convenient for training purposes. All images were collected from the BIRDS 400 dataset found at https://www.kaggle.com/gpiosenka/100-bird-species. One important thing to note is that 85% of the images are from the male of the species and only 15% are female, this is due to the male having more colors and features that make them more easily recognizible. For training purposes it makes

it more convenient, but this imbalance in our data could present a problem if trying to classify the female bird of the species.

More experiments need to be performed to see the possibility to use as much data for building our models without sacrificing performance. Due to the heavy load of dealing with so many pictures, some measure might need to be adopted when building our models for prototyping purposes. These might be randomly taking a smaller sample of the pictures and/or reducing the number of species that can be identified, in order to make the classifier development more manageble. This will allow us to focus more on the actual implementation and comparison of each approach.

3   Methodology

Data pre-processing will be done before building any of the models by downsampling the images, selecting only a couple of species to classify, and taking a random sample of images from each species. This will be done for prototyping purposes and to focus more on the actual building of each of our models. Once each implementation is validated and tested more data will be included in the overrall training and testing step of each approach. The final results will be presented in the evaluation and final results section. Below we give a quick glance to each proposed methodology to be used for building our image classifier. Furthermore, in-depth analysis of each approach, mathematical explanations, and visual graphs of results; will be provided in the final report.

3.1   Dimensionality Reduction

For our dimensionality reduction approach we will use principal component analysis to find the most representative features combinations for each species of bird. We will do this by finding the first n principal components of each species we want to classify, and calculating the projection residual between the eigenspecies and a new image y. The lower the residual, the more likely it is that the new bird y belongs to species k. A threshold could be implemented as means of determining if a bird is similar enough to be considered any species k. We will evaluate the performance of our model by predicting a number of different test images and looking at different classification performance metrics such as precision, recall, F score, and more. Finally, advantages and disadvantages of this approach will be discussed. This implementation is pretty similar to the eigenfaces example presented in class.

3.2   Density Estimation

A different approach for creating our bird classifier could be using density estimation. If we were to divide all of our training pictures into n different clusters, and group together all the images that look similar enough. We could then use a new input image x and classify it calculating the lowest residual between that image and all the different centroids. Each cluster will be assigned the majority label of that specific cluster as its label. Different measures would need to be taken into account to make sure the model performs well. Things like the homogeneity of each cluster and the overrall accuracy of such a model. More experiments need to be performed to determine if a parametric or non-parametric approach would be best in this case for performing clustering. The advantage of this method is that the number of clusters doesn't need to match the number of species, what would matter most would be the homogeneity of each different cluster with respect to the labels of each training image. This would give us the best opportunity for generalization and high predictive accuracy for future tests.

3.3 Classification

In this step we will be building a number of different models using traditional classification methods. KNN, SVM, Kernel SVM, Logistic Regression, and others will be used to create a classifier for our bird species identifier. We will compare them all and discuss the disadvantages and advantages of using each. Classification and model specific metrics will be presented for technique evaluation, and seeing which one is more efficient for the task. Some of this algorithms might be computationally expensive to use due to the big number of data, if this is the case the results will be presented as such and given as a suboptimal solution to our problem. More experiments will be done regarding this case.

3.4 Deep Learning

Deep learning is a subfield of machine learning that specializes on reproducing the way the human brain thinks to solve certain problems. It does this by using artificial neural networks; a mathematical representation of a composition of neurons signaling to each other. This algorithm has proven to have an endless number of real life applications when it comes to using data to gain insight on an input. In this case, we will implement a variation of this approach called a convolutional neural network. The specific inner workings of this type of network will be discussed and we will do an implementation applied to our training data using the keras library in python. We will use the dropout method to prevent overfitting, and validation to find the optimal parameters for our model. We will use classification metrics to evaluate the quality of our algorithm and the time it takes to produce an output as well.

4 Evaluation and Final Results

In this final section the most meaningful metrics from each one of the 4 applied techniques will be presented and compared. Further discussion on the performance of each and further aplications will be mentioned. Testing with a higher number of data will also be performed at this stage to see how the best technique generalizes when presented with a more real world example. Using our better perfoming approach, we will build our bird species identifier and give the final evaluation metrics of it's implementation. If good results are obtained from initial experiments and no big issues arise while developing each one of the models, a video classifier that takes one frame every j frames and which outputs a label if a bird appears in one of them, might be a good real life implementation of our optimal approach. This will only be done if the scope of this project doesn't prove to be too heavy on one person. If not, scalibility and further possibilities will also be discussed in the conclusion section.