Kimberly Adams
DSC 680 Spring 2024

Bird Photo Identification

Bird watching is the enjoyment of watching of watching wild birds. It ranges a full spectrum of activities from simply watching, photographing, identifying, and/or feeding birds to the avid birdwatchers who may travel to see unique birds or "collect" them on their life list of sightings. The nice thing is that you don't need any equipment for basic birdwatching except your eyes and/or ears. This means it is accessible to just about everyone since birds are everywhere.

Birdwatching is a popular hobby with a large economic impact. According to the US Fish and Wildlife Service 2022 National Survey of Fishing, Hunting, and Wildlife-Associated Recreation, more than 148.3 million people watched wildlife and contributed nearly $250.2 billion to the U.S. economy in 2022. Birders contribute to the economy by spending on equipment (e.g., binoculars, field guides, feeders, and baths) and travel-related expenses (e.g., gas, accommodations, and park entrance fees).

In addition to an economic benefit, contact with nature also has health benefits for the participant. Many studies have shown that being in nature can help improve a person's mood by reducing feelings of isolation and creating a sense of connection to something greater. Other studies found reduced depression, anxiety, stress, fatigue, tension, and chronic pain as well as increased attention spans. Getting outside is also beneficial in general by increasing the body's uptake of vitamin D from sunlight exposure.

If we can develop tools to foster a love of birdwatching for those just starting, we can provide many benefits to both the individual and society. To that end, for this project, I attempted to make an image classification system that could identify birds in photos. By aiding in identification, the tool would be able to increase the connection between the beginning birder and the birds they see.

Data

When looking for data for this project, I found three different potential datasets on Kaggle. I decided to go with the smallest dataset species-wise with the hope that this would simplify the model (https://www.kaggle.com/datasets/umairshahpirzada/birds-20-species-image-classification). This dataset includes 20 species in it with 3,208 training images, 100 validation images, and 100 test images. The species range from small songbirds to waterfowl and shorebirds to parrots and raptors. It also includes species from multiple locations.

Each species in the training subset has between 133 to 187 different images (Figure 1). The images vary in terms of background scenery, lighting, and presentation of the individual bird such as posture, angle, and direction facing (Figure 2). All of these characteristics aid in mimicking the variation that would likely be included from random observations.

The images include both males and females as well as juveniles. There are more images of males than females and juveniles though which will make the model less able to distinguish non-male individuals. Since males tend to be more showy and easier to differentiate and females and juveniles less so, these are challenging even for humans to distinguish between. This will be a limitation to any model made with this dataset.

The test subset contains similar variation but only contains five images for each species (Figure 3). Same with the validation subset (Figure 4). Finally, there is one more subset containing six total images to predict. These six images contain five African Crowned Cranes and one Bald Eagle image.

All images were assumed to be correctly identified within the dataset. Some images for familiar species were spot-checked to verify labeling.

Methods

After general exploration of the data, the images were fed through the Keras ImageDataGenerator to provide some preprocessing and standardization of the training images in terms of things like brightness, scale, and rotation. While only the training images were augmented with additional standardization effects, all images were rescaled and resized.

A Convolutional Neural Network was used to create a sequential model with multiple convolutional, max pooling, and connected layers. ReLu and softmax activation functions and dropouts are also included within the layers. Each layer adds a further level of analysis at different scales and focuses. What one layer misses, hopefully another will pick up on while other functions help keep the model from getting overly complex. Overall the model has twelve layers of analysis.

The compiled model was fed the training images and attempts to categorize the bird images into similarities for each species. The fit was performed in 200 image batches for 20 epochs. The trained model was then applied to the test images and the accuracy assessed. This took some time and would presumably take longer if the number of species was increased.

Analysis

With each epoch repetition of the model fit, the accuracy of the model increased while the loss decreased showing slow but steady improvement until it started to flatten out and the iterations were stopped (Figure 5). The model accuracy for the training images peaked at around 60%, for the validation images at 69%, and for the test images at around 78% (Figure 6).

Looking at the confusion matrix for the test data, most species were correctly identified, but the model struggled with certain species more than others (Figure 7). Some of the species the model struggled with the most included the African Pied Hornbill and Albert's Towhee followed by the Abbott's Babbler, Abyssinian Ground Hornbill, African Oyster Catcher, Alpine Chough, and American Avocet.

When the model was applied to the six designated images for prediction, it successfully labeled the five African Crowned Cranes but failed at correctly identifying the Bald Eagle which it labeled an Abbot's Booby (Figure 8). This is actually to be somewhat expected as the model had not been trained to recognize Bald Eagles and thus labeled it as the closed match to what it knew which in this case was a similarly colored white-headed bird with a bright yellow beak. This once again indicates that there is some merit to the model, but it is not perfect and that it only works to recognize the species it has been trained on.

Conclusion

The fact that the model did better with the test images than it did with the training images indicates that it was successful. This is echoed in the high 78% accuracy rate for the test images. However, the goal was to create a model that had an accuracy rate of 90% or better and this model did not live up to that criteria. There may be room for further adjustment of the model input parameters within the layers that comprise it by adding additional layers or tweaking the existing ones.

There are a couple of bird identification phone applications that are already on the market. Most are strictly digital versions of field guides, but some, like the Cornell Lab of Ornithology's Merlin app, attempt to provide identification via both photo and sound samples.

Having used the Merlin app many times in my own birdwatching, I can tell you that it too is nowhere near perfect. One article I found suggested it has an accuracy rate of 90% for photo ID and another said 90% for sound identifications (depending on a high-quality input sample). Both of these rates are better than the model created for this project but still have flaws. The way the app makes up for it though is by both providing multiple potential matches and also providing sample pictures or sound bites of each match so the user can verify for themselves that the bird is what they saw or heard. It also includes a place to submit feedback for bad matches to aid in continual improvement.

Both the model created here and apps like Merlin suggest there is some promise in creating machine learning models for animal identification. Although the models are not perfect, models such as that used in the Merlin app are good tools to help beginner or intermediate birdwatchers advance their skills. Sometimes being presented with the incorrect match can be an opportunity to hone observation skills as to why that match is incorrect and encourage the user to keep looking for the correct one.

In the end, the important thing is to connect people with the nature around them. Tools like this might not be everyone's cup of tea, but if it encourages a few new people to get outside and actually look around, it may inspire a love for the wild world around us before it is all gone. It might also encourage users to develop not only a love for birdwatching but perhaps photography as well. Both effects benefit the individual with increased skills and an improved mental state but also benefit the economy as these folks travel around to see and photograph new birds beyond their backyards. Money talks and the more people spend to go see nature and natural places, the more those places are likely to be protected and preserved for future generations to enjoy and benefit from.

Ethical Considerations

Assuming the pictures contained in the datasets were freely available and unlicensed, there should not be any ethical concerns with this project. The description of the dataset says the images were collected through Google searches and then processed via cropping and other techniques.

A misidentified bird should not cause any harmful effects to anyone using the model and should not cause harm to the bird either. The most likely impact would be the user posting an incorrectly labeled photo and a different birder correcting the identification.

There is a possibility if the model moved into production such that users could submit their own photos that more people might be willing to get too close to birds to get a good photo. This would be considered disturbing and harassing the bird and is illegal in many places. The model may be limited to shots of a certain quality which might in turn limit users to those with strong telephoto cameras and decent photography skills which in turn might be unlikely for a beginner birdwatcher.

Finally, there is some small risk that a tool like this that makes identification "too easy" might discourage beginning birders from learning the field markings that distinguish each species. Why learn it when you can just plug it in and get the answer? I personally doubt this would be the case for most people as through simple repetition the user would start learning patterns. For example, if I saw a red bird in my yard it is probably the same Northern Cardinal the app identified for me yesterday which the blue bird is a Blue Jay. At some point, entering the photo of a familiar species into the app would be more cumbersome than personally identifying it. In the end, the biggest thing is getting people outside, and even if they don't utilize the tools to learn and build their skills, at least they are outside interacting with nature.

Appendix

Figures:

Figure 1. Images in the Training Dataset by Bird Species

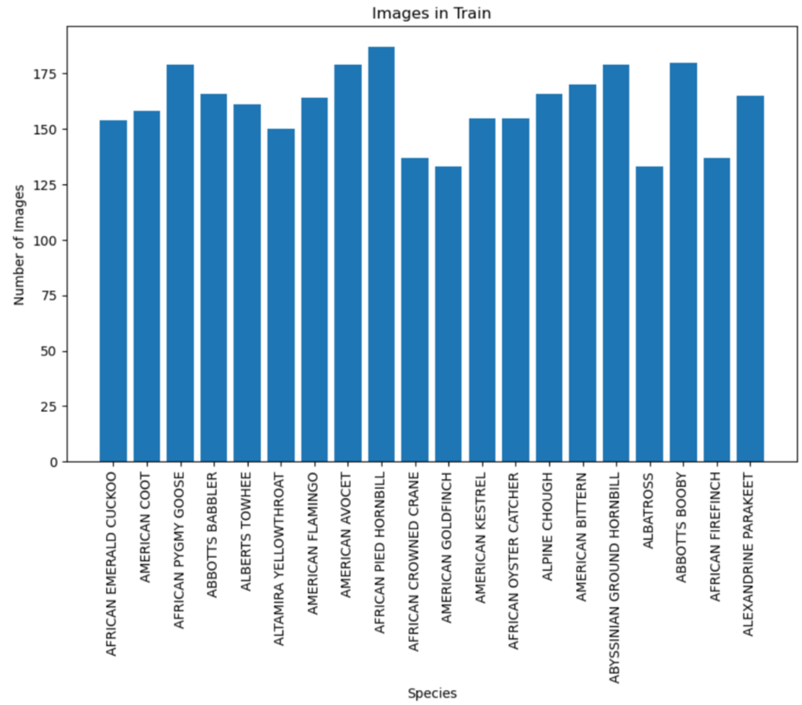

Figure 2. Random Sample Images from the Training Data.

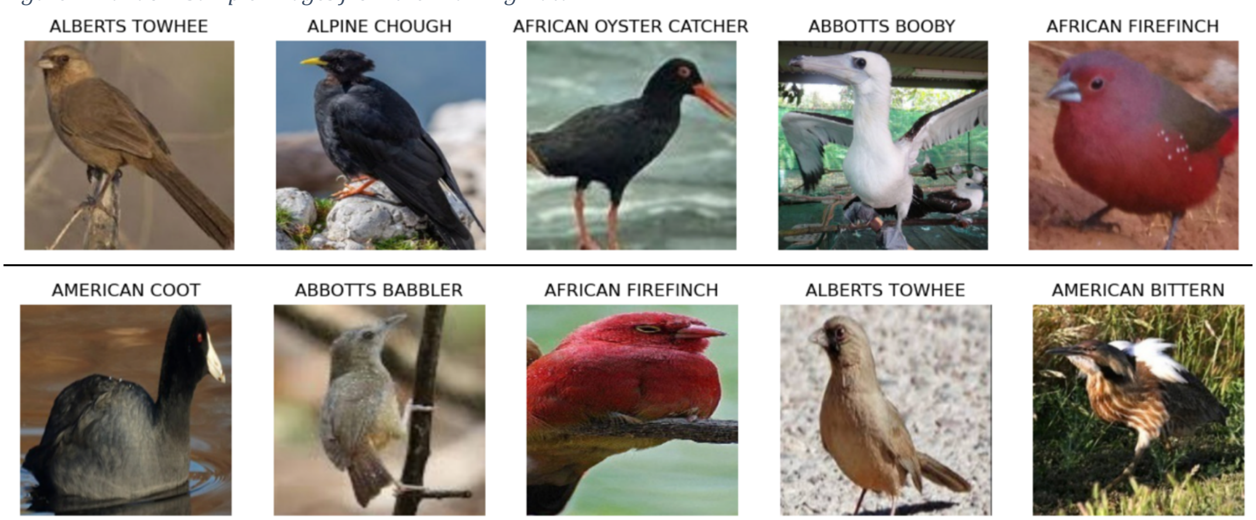

Figure 3. Images in the Test Dataset by Bird Species

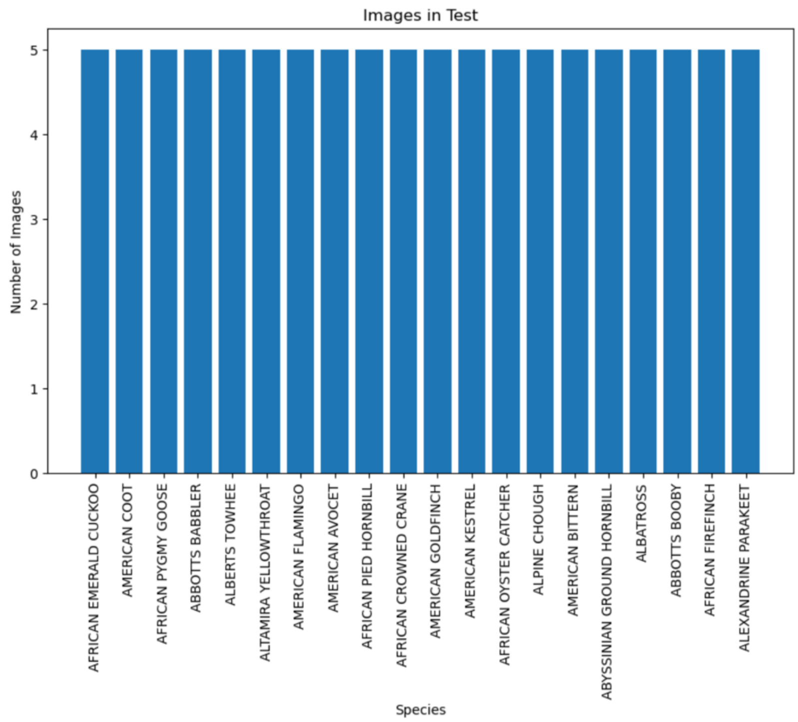

Figure 4. Images in the Validation Dataset by Bird Species

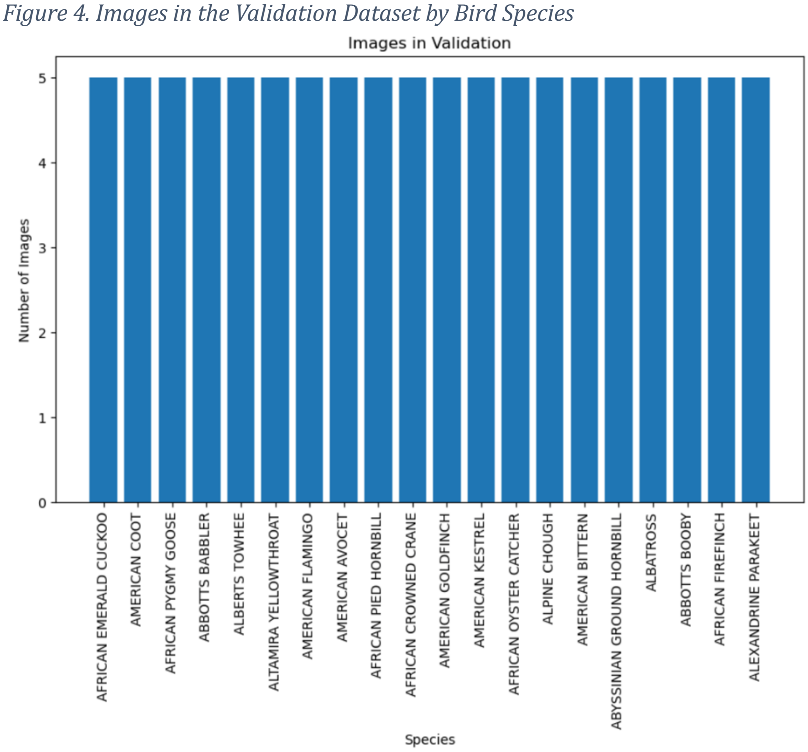

Figure 5. Model Training Accuracy and Loss Over Training Epochs

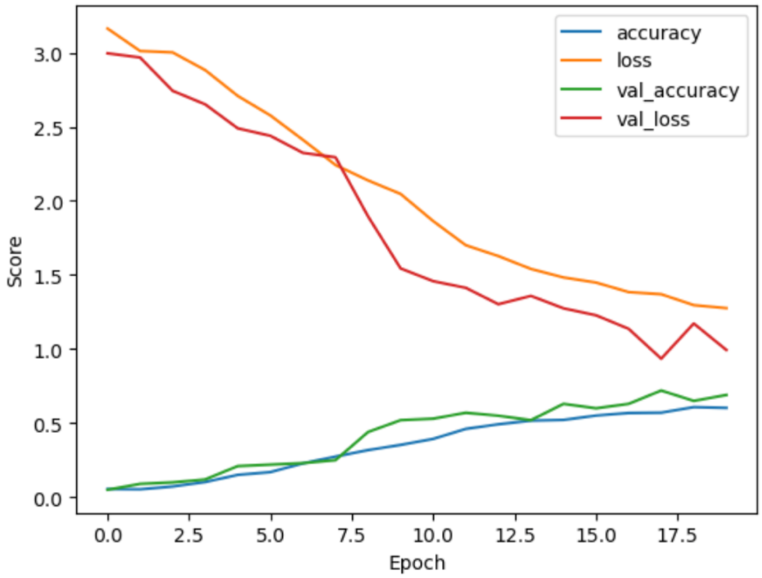

Figure 6. Model Accuracy on Different Data Subsets

Dataset	Accuracy	Loss
Training	0.599439	1.24583
Validation	0.69	0.993831
Test	0.78	0.745995

Figure 7. Confusion Matrix for Model Applied to Test Images

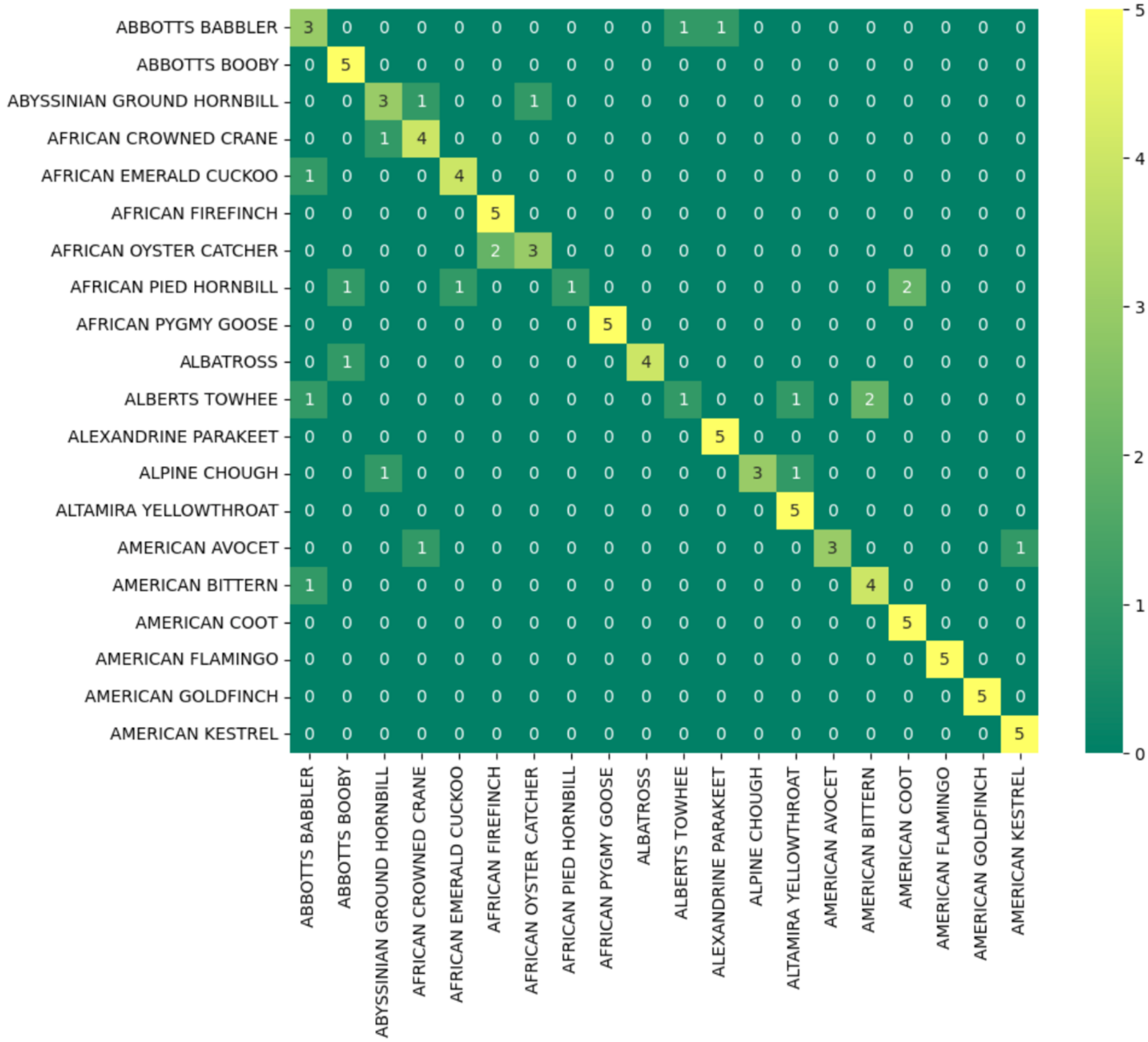

Figure 8. Designated Images to Predict with Predicted Species ID.

References:

2022 National Survey of Fishing, Hunting, and Wildlife-Associated Recreation - Documents - USFWS National Digital Library. (n.d.). https://digitalmedia.fws.gov/digital/collection/document/id/2321/rec/1

Adams, J. U. (2023, February 21). Birding With Benefits: How nature Improves our mental mindsets. *Audubon*. https://www.audubon.org/magazine/winter-2019/birding-benefits-how-nature-improves-our-mental

Analytics Vidhya. (2024, February 18). *Image classification using CNN (Convolutional Neural Networks).* https://www.analyticsvidhya.com/blog/2020/02/learn-image-classification-cnn-convolutional-neural-networks-3-datasets/

Barreladmin. (2019, May 24). *Cornell Tech - Merlin Bird Photo ID Mobile app launches*. Cornell Tech. https://tech.cornell.edu/news/merlin-bird-photo-id-mobile-app-launches/

Birding in the US Report_2016_508 accessible; - Documents - USFWS National Digital Library. (n.d.). https://digitalmedia.fws.gov/digital/collection/document/id/2252/

Can birdwatching really improve our mental health? The science says yes. (2021, October 15). BirdLife International. https://www.birdlife.org/news/2021/10/05/can-birdwatching-really-improve-our-mental-health-the-science-says-yes/

Cox, D. T. C., Shanahan, D. F., Hudson, H. L., Plummer, K. E., Siriwardena, G. M., Fuller, R. A., Anderson, K., Hancock, S., & Gaston, K. J. (2017). Doses of Neighborhood Nature: The Benefits for Mental Health of Living with Nature. *BioScience/Bioscience*, biw173. https://doi.org/10.1093/biosci/biw173

Indian-Birds-Species-Image-Classification. (2023, March 21). https://www.kaggle.com/datasets/ichhadhari/indian-birds

BIRDS 20 SPECIES- IMAGE CLASSIFICATION. (2023, April 5). https://www.kaggle.com/datasets/umairshahpirzada/birds-20-species-image-classification

BIRDS 525 SPECIES- IMAGE CLASSIFICATION. (2023, April 20). https://www.kaggle.com/datasets/gpiosenka/100-bird-species

Ratcliffe, E., & Gatersleben, B. (2013). Bird sounds and their contributions to perceived attention restoration and stress recovery. *Journal of Environmental Psychology*, *36*, 221–228. https://doi.org/10.1016/j.jenvp.2013.08.004

Tiller, E. R. (2024, April 11). Boost your mental health with bird-watching. *Birds and Blooms*. https://www.birdsandblooms.com/birding/birding-basics/birding-health/

Wilson, Herb. (2022, September 11). *Merlin and the Use of Artificial Intelligence in Bird Identification. Maine Birds.* https://web.colby.edu/mainebirds/2022/09/11/merlin-and-the-use-of-artificial-intelligence-in-bird-identification/#:~:text=Merlin%20will%20then%20identify%20the,of%20the%20sound%20identifications%20right.