Investigating the Use of Artificial Intelligence to Process Camera Trap Data

Kassidy Robinson[1], Jennifer Higbie[2],

[1]Environmental Studies, Binghamton University, Vestal, NY 13850
[2]Environmental Protection Division, Brookhaven National Laboratory, Upton, NY 11973

ABSTRACT:

Wildlife camera traps are an invaluable tool in wildlife biology. The use of wildlife cameras has many benefits, including saving time in the field and gaining a higher quantity of data, but also has a major drawback. This wealth of data is meaningless if it cannot be efficiently sorted and processed. Camera traps often yield an overwhelming number of photos. The goal of this project was to attempt the implementation of artificial intelligence (AI) programs that could make it possible to process the large quantity of wildlife data collected by Brookhaven National Laboratory. The first approach was to implement a free package for R-studio called MLWIC (Machine Learning for Wildlife Image Classification). After challenges, a second alternative was considered called Wildlife Insights Beta. These attempts resulted in the findings that: 1. MLWIC may not be compatible with an updated Windows computer or alternatively that a computer science/coding expert is needed to determine if the program is viable in the future; 2. Wildlife Insights is a useful starting off point which reduces the amount of labor for the user by effectively sorting out blank photos, but is still in Beta development and will require more AI training in the future in order to achieve high enough accuracy that would warrant consistent use by the laboratory. I have processed 12,426 photos and provided usable wildlife data for potential population estimates for various wildlife species present at the laboratory. After testing these artificial intelligence programs, the laboratory may use this information to determine whether it is warranted to invest funds towards an artificial intelligence program that is more accurate and user friendly than the previously tested free programs. As a result of my internship, I have gained experience in coding and in working with and troubleshooting artificial intelligence.

INTRODUCTION:

Since the rise in use of wildlife cameras by biologists in the late 20th century, researchers have been able to uncover unique footage of wildlife. These camera traps enable biologists to non-invasively monitor wildlife populations by setting them on a timer and/or using motion sensing technology. This innovation saves researchers time in the field and legwork of observing at unfavorable hours of the night, as well as increases accuracy of encounters by decreasing the disturbance to the wildlife which captures more authentic behavior without the presence of humans. Additionally, camera traps have illustrated higher accuracy than observed in traditional road sampling methods with respect to an endangered subspecies of white-tailed deer because doing so can reduce sampling bias associated with viewing deer at their habitat edges [1]. This is of particular

Top left and bottom left: white-tailed deer, *Odocoileus virginianus*. Top right: red fox, *Vulpes vulpes*. Photos captured in the wildlife survey at BNL taking place from January 19 – March 3.

interest to Brookhaven National Laboratory (BNL) because each year researchers conduct a road sampling of white-tail deer but have recently encountered numbers they believe to be an underestimation. Utilizing widespread camera trap data, researchers may find alternative methods to develop accurate population estimates for an ecologically significant species.

Further, the laboratory has populations of red fox and historic sightings of grey fox, but their den locations are majorly unknown, and their behavior is elusive. It is difficult to spot foxes in field surveys, but camera traps can remedy this. Camera trap data collected on red foxes has been used to estimate abundance and population density by identifying individuals based on physical characteristic [2] . This data in turn can be used to monitor the prevalence of diseases such as mange that can inhibit the success of their populations as well be used to track where foxes are seen most frequently and observe where they are seen eating to hone in on possible locations of dens [3], which benefits efforts to monitor kit populations and success. This method could be utilized by BNL to further enrich their monitoring efforts of red and gray foxes on campus if camera trap footage was more readily processed.

However useful these cameras have proven and have the potential to be, they result in a massive quantity of data, many of which do not contain any wildlife. This data is beneficial for gaining a deeper understanding of wildlife but is often not analyzed quickly due to the amount of data collected. Going through the data by hand is time consuming and can possibly cause physical harm to the researchers over time in the form of eye and wrist strain from the incessant clicking through photos. Without an efficient method of processing camera trap footage, population models and estimates may not be developed, and highly valuable wildlife data will not be put to any significant use. This project aims to find an appropriate solution to hand sorting photos collected by BNL's camera traps and to test the chosen method on photos gathered so that in the foreseeable future the lab can employ one of the aforementioned methods to study local wildlife populations.

Camera traps associated with Brookhaven's four-poster tick management systems project generate roughly one million photos per year. It is the goal of the Environmental Protection Division to implement widespread camera traps at BNL, which would generate hundreds of thousands, if not millions, more per year. It is clear that the laboratory is in need of a better system than hand sorting photos. If the laboratory is able to process photo data consistently and efficiently, camera traps could be established and utilized for year-round population monitoring. This goes for many significant species on the campus, including deer, gray foxes, red foxes, turkeys, and in the future, possibly coyotes as they expand their range and inevitably become more abundant on Long Island. This initiative would greatly benefit wildlife management on campus and provide opportunities for novel research.

OBJECTIVE:

The overarching objective of this project was to work with artificial intelligence (AI) to determine a method suitable to the laboratory's needs to sort wildlife photo data, to use this system to sort 12,426 new images of photo data, and to assess the usefulness of the artificial intelligence programs used.

METHODS:

Part 1. Experimental Setup

On February 19, 2021, 18 camera traps were established across BNL's undeveloped areas in a uniform grid that maintained a minimum 0.5 km distance and left to collect data until March 3, 2021. They were attached to trees and were set to photograph when motion was detected. The Moultrie M50i cameras utilized infrared flash as not to disturb or scare wildlife away. Scent traps were utilized in order to attract carnivores.

On March 30, 2021, 14 cameras were set out for a second survey. This survey used the same methods as the first survey. Camera cards were pulled on April 9, 2021.

Part 2. MLWIC

My work began with the attempt to pick up where a previous intern, Coral Salort, left off, testing the viability of the MLWIC (Machine Learning for Wildlife Image Classification) package in R-Studio [4]. To clarify, the program was no longer installed so I began by reinstalling the software by following her detailed instructions [5].

MLWIC [6][7] was created by Mikey Tabak. It is an open-source program available on GitHub for download as a free package into R-studio that was designed to evaluate wildlife images using a species level model as well as train the program to meet the user's classification needs [4]. The model utilizes the TensorFlow framework in which a convolutional neural network with ResNet-18 architecture was trained to sort up to 2,000 photos to the species level per minute with accuracy reported up to 97.6% [4].

Salort's work outlined steps for installation and use both for Mac and Windows [5], although it is important to note that the program was designed for use on Mac and less feedback and instruction have been offered for Windows. The process I followed for implementing MLWIC is briefly outlined below:

Installation and Setup

1. Downloaded R/R-Studio
2. Downloaded Anaconda
    - Set up environments in Anaconda
3. Download test folder L1 and example folder from GitHub
4. Setting up in R-Studio

- Set conda_path and wdir (location of anaconda and working directory, respectively), and install the "devtools" package.

Function Classify(): This is the first function within MLWIC that must be used in order to sort photos. It accesses a csv file with names of photos and a file with photos themselves and attempts to output them in another designated location sorted by species.

1. Attempted to run the code. Received several different error messages, and I suspected I may have done something wrong in installation. I uninstalled anaconda and r and reinstalled everything within carefully according to directions from Salort's appendices and the GitHub guide.
2. Ran the code again. This time no error messages. No results or progress either. Could not complete the next step.

There are several ways I dealt with challenges faced when implementing this program. Firstly, to overcome my lack of coding background, I spent one week completing a course on Codecademy[8] to ensure I knew the basics of coding in R-studio. I also compiled a list of references and cheat sheets for coding in R to reference if I was having trouble remembering a function or how to interpret the syntax of a line of code. Next, when attempting to implement MLWIC, I cross referenced instructions by Mikey Tabak[6][7] and supplemental instructions by Coral Salort[5]. This step was crucial because there are not many sources available on MLWIC since it is relatively niche and was released in 2018. When facing challenges running the Classify() function in R-Studio for MLWIC, I posted my question with an error code I was receiving on the MLWIC GitHub page. The creator, Mikey Tabak, along with other users and collaborators have responded to questions on that page in the past, but I received no answer. After not having any way to move forward, I uninstalled and reinstalled R-Studio, Anaconda[9], the L1-Test folder, the example folder, and then MLWIC within R-studio. This was done to ensure that every associated program was installed in the correct version that was required for MLWIC to function. After doing so, I attempted to run Classify() and this time, not only did it not produce results, but it did not provide any error code that I could troubleshoot or work off of.

Part 3. Wildlife Insights

After reaching a standstill with MLWIC in R-Studio, I found Wildlife Insights Beta [10] as an alternative. It is also free and open source like MLWIC, however the technology is available on their website to use rather than to download. Further, Wildlife Insights does not require coding or computer science skills to utilize whatsoever, making it much more accessible. Some of the criteria I reviewed in my preliminary research on Wildlife Insights is shown below in figure 1.

Species	Common name	Number of photos in database	Percent of database	Precision	Recall
Odocoileus virginianus	White-tailed deer	10,511	0.0000001%	94.81%	50.60%
Canis latrans	Coyote	41,544	0.0000356%	68.75%	64.88%
Canidae urocyon cinereoargenteus	Gray fox	18,378	0.0000158%	90.47%	87.28%
Vulpes vulpes	Red fox	1,735	0.0000015%	90.62%	41.43%
Meleagris gallopavo	Wild turkey	4,181	0.0000036%	100%	65.48%
Procyon lotor	Northern raccoon	57,059	0.0000489%	90.60%	76.56%
Sciurus carolinensis	Eastern gray squirrel	29,557	0.0000253%	96.66%	90.61%
Homo sapiens	Human	290,190	0.0002488%	82.32%	84.55%
Felis catus	Domestic cat	5,411	0.0000046%	needs more data	needs more data
-	Blank	4,281,353	0.0036705%	97.36%	46.84%

Figure 1. Select statistics provided by wildlifeinsights.org for several species of interest that were likely to be captured by BNL's camera traps. (Data obtained from wildlifeinsights.org)

Application and usage methods are outlined below:

1. Creating an account.

2. Becoming a Trusted Tester of the Beta program
    a. This involved emailing back and forth with one of their representatives at info@wildlifeinsights.org and filling out an online form with details of the project.
3. Optional: taking the informational class.
    a. I completed this step after having used the program to sort most of my photos. Info@wildlifeinsights.org was available to answer my questions before then. I went to the instructional seminar to gather information about the development of Wildlife Insights' features that might prove useful in the future.
4. Creating a titled project within Wildlife Insights.
5. Uploading the photos to the project folder.
    a. Create sites with information about the camera traps.
    b. Upload photos. It works best in batches of up to 500. If a site has more, it may be best to split it up.
    c. Review suggestions. Accept suggestions, mark as blank, or add/change a species seen.
    d. Repeat for all sites and photos.
    e. View already sorted photos in "categorized" and stats in "summary".
6. Offer feedback to the Trusted Tester Team to assist in development of the beta phase.
    a. Filled out request feature form.
7. Request that photos are utilized in retraining the model (approved).
    a. Modify classification of all "blank" photos into specific categories (true blank, human, dog, car) so they can be utilized for AI training.

RESULTS:

Part 1. MLWIC

MLWIC was not able to be used to analyze any photos.

Part 2. Wildlife Insights

Wildlife insights was effective for a small sample size of photos. I uploaded and categorized about 10,000 photos from over a three-month period on their website over a span of about three weeks of work. However, it had limited effectiveness and is likely limited to these small sample sizes due to its inaccuracy which requires the user to double check and change many of its suggestions.

Part 3. Camera Trap Results

Wildlife encounters as recorded from the first survey, taking place from January 19 – March 3, 2021, are displayed below in figure 2.

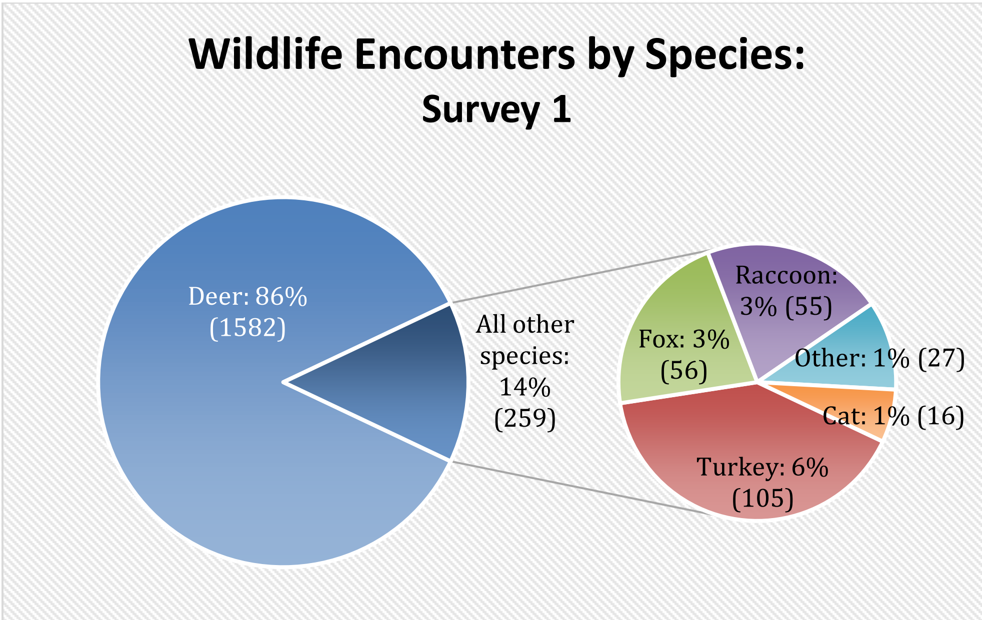

Figure 2. Wildlife Encounters by Species, Survey 1. Shown above is the number and percentage of all wildlife species seen from January 19 – March 3, 2021 on the camera trap footage. Omitted from the figure is the proportion of 7327 blank photos which comprise 79.92% of all photo data collected from this survey. Shown on the left is the breakdown of the remaining 20.08% which contained wildlife.

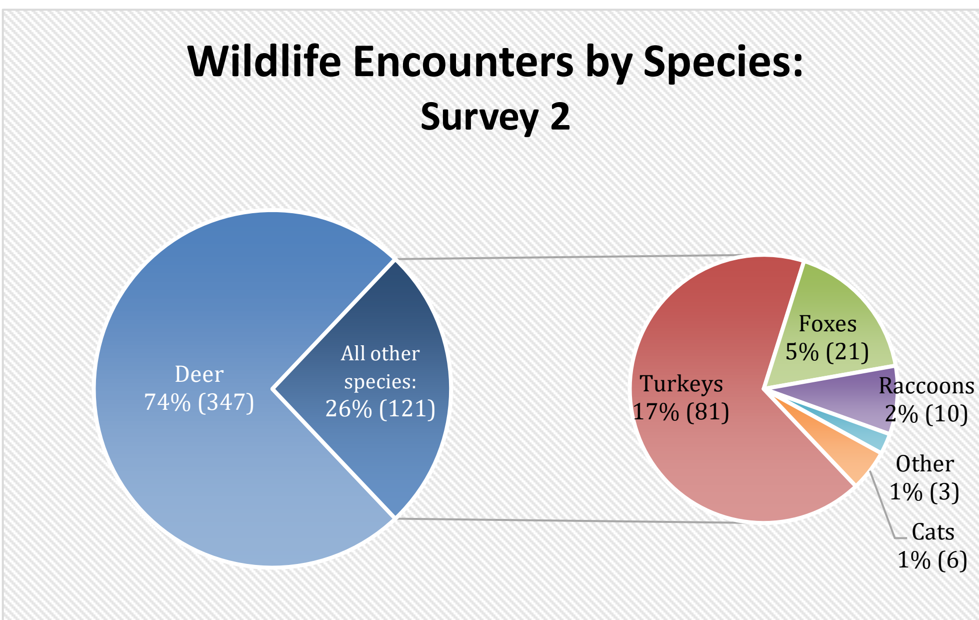

Figure 3. Wildlife Encounters by Species, Survey 2. Shown above is the number and percentage of all wildlife species seen from March 30 – April 9, 2021 on the camera trap footage. Omitted from the figure is the proportion of 2,798 blank photos which comprise 85.88% of all photo data collected from this survey. Shown above is the breakdown of the remaining 14.12% which contained wildlife.

DISCUSSION:

Part 1. MLWIC

The current prediction on the failure to implement MLWIC is that it is only compatible with select versions of Windows. It is possible that Windows has updated and has rendered this software incompatible. If this is the case, this program may not be compatible for the needs of BNL, in that their federal computer systems must be kept up to date.

It is very difficult to troubleshoot any MLWIC problems because the package was recently published in 2018 and it was not created for Windows. There are few resources available to learn more or to fix these issues. I posted on the online forum page of GitHub, but no one had responded. There are no instructional videos on installation or use of the package because it is new and very niche in terms of computational programs.

The attempt at running this program has limitations, however. I have had no previous background in computer science, coding, artificial intelligence, or R-Studio, therefore it is possible that the error lies in my implementation methods, not in the program or its compatibility with Windows. It is possible that an expert in computer science or artificial intelligence programming would be able to remedy these issues, however.

Part 2. Wildlife Insights

Wildlife Insights was able to process all of the photos, but not without some technical difficulties and inaccuracies. Its interface is incredibly user friendly, and its Computer Vision (CV) efficiently sorts out a majority of the blank photos, which saves a significant amount of manual work for the user. However, the CV consistently mistook some species for others, lacking the ability to take into account location of survey in its species suggestions, as seen in figure 4.

Computer Vision suggestion	True species ID
Mule deer	White-tailed deer
Coyote	White-tailed deer
Coyote	Fox
Panther	Domestic cat
Lynx	Domestic cat

Figure 4 (Left). Differences between Wildlife Insight's Computer Vision suggestions as compared to true species identification as determined while reviewing survey photos.

Figure 5 (Below). Model performance of frequently suggested species as compared to true species. As one can see by viewing figure 4 and figure 5 side by side, the overarching reason for the Computer Vision's faulty suggestions is due to the abundance of other species not found in our survey region. The CV is only able to use what its data has been input by other users, so if there are 215,928 photos of mule deer in the database and only 10,511 photos of white-tailed deer, it may assume false scarcity of white-tailed deer or not have enough data yet to correctly distinguish them from mule deer. This same logic can be applied to the other false suggestions as well. (Data obtained from wildlifeinsights.org)

Species	Common name	Number of photos in database	Percent of database	Precision	Recall
Odocoileus hemionus	Mule deer	215,928	0.0001851%	95.72%	91.35%
Odocoileus virginianus	White-tailed deer	10,511	0.000000091%	94.81%	50.60%
Lynx rufus	Bobcat	34,862	0.002990%	96.04%	82.37%
Leopardus pardalis	Ocelot	21,024	0.00180%	95.25%	64.79%
Puma concolor	Puma	20,666	0.001770%	93.52%	80.36%
Felis catus	Domestic cat	5,411	0.0000046%	Needs more data	Needs more data
Canis latrans	Coyote	41,544	0.0000356%	68.75%	64.88%

This program is still in its beta phase and its AI lacks accuracy for identification of species found in the Northeast United States. This low accuracy results from a lack of users in this region, therefore the CV model has not been trained sufficiently with species from this area as compared to other species in the United States, as illustrated in figure 5. To increase accuracy of the identification of wildlife species of interest to Brookhaven, there are two options: 1. Continue using the model, correcting the suggestions by hand and 2. Request that your data is used to retrain the model. I was recently approved for the data to be used to retrain the model. In the future, this method can be used to increase region specific accuracy of the model during a testing period followed by a projected increase in accuracy of the model's predictions. This increase in accuracy will result in saving time going back and sorting by hand after uploading the photos to correct false suggestions.

Another possible avenue to increase this program's accuracy could be achieved through requesting a feature. I used their online form to suggest implementing a feature, as previously mentioned, that I believed would benefit users and CV accuracy. The suggestion was to incorporate information about location of the field surveys into AI decision making. Meaning, by telling the AI what region the data was collected from before it sorts the photos, it could hypothetically access a database of what type of wildlife are likely to be seen there so it can make more informed suggestions. This would remedy the problem I encountered of having white-tailed deer consistently identified as mule deer. This method of requesting features to be added is more so a long-term solution for improving Wildlife Insights overtime, because any major feature would take some time to be developed. However, the more researchers invest their time and energy into using this system, the more accurate and useful it becomes to themselves and others.

Overall, Wildlife Insights was helpful in processing this small batch of photos of about 10,000 from over a three-month period. It reduced time taken and strain on the user manually sorting the photos by distinguishing non-animal photos with very high effectiveness. However, the goal of the Environmental Protection Division at BNL is to be able to monitor wildlife across the campus year-round, therefore another method may be more realistic to accommodate for the high volume of photos they may have. In the future, Wildlife Insights developers plan to implement several new features that will improve performance and usefulness. These include data download capability for the public, further AI improvements every two months, instantaneous analytics for user's data, and data management and processing. Along with these upcoming features and the potential for improvements in local species accuracy, Wildlife Insights could prove to be an asset to the laboratory and provide a long-term free solution to arduously hand-sorting hundreds of thousands of wildlife photos per year. This is especially true if the researchers were to upload the wealth of white-tailed deer photos, they already collected from the four-poster project onto Wildlife Insights. Doing so would provide plenty of data to increase accuracy for white-tailed deer identification and potentially increase the Wildlife Insight's usefulness for future surveys.

Part 3. Data Collection

It is important to note a few limitations and shortcomings in the data collection and experimental setup phase of this survey that may have affected its results. The camera trap locations utilized scent traps with the goal of attracting carnivores to the sites. There were several major snowfalls during the first survey and rainfall during the second survey which can be verified by the National Weather Service's records [11], that influenced the scent traps and undoubtedly decreased their effectiveness. Therefore, it can be generalized that the survey photos received after the first major snowfall and rainfall during each survey were random encounters likely not influenced by the presence of bait, but further analysis is necessary to say definitively.

NEXT STEPS:

Looking forward, this data has potential to be used for several research initiatives. BNL contacted the New York State Mammal Survey to contribute photos to populate wildlife species on Long Island. Meaning, BNL will submit their photos so that other users and researchers can view data on what kinds of mammals are in the area and assist in the first statewide mammal survey in New York since 1971. This is likely going to be an ongoing project for the department and provide further research opportunity. In the future, it is the hope of BNL that other local agencies and organizations from Long Island will join the efforts of contributing to the New York State Mammal Survey to help construct a more comprehensive picture of wildlife populations on Long Island in order to better manage and conserve the local fauna.

ACKNOWLEDGEMENTS:

This project was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI). I would also like to thank my mentor, Jennifer Higbie, and the other staff at Brookhaven that made this project possible. Thank you, Jennifer Higbie, Timothy Green, and Kathy Schwager.

REFERENCES:

[1] Roberts, C.W., Pierce, B.L., Braden, A.W., Lopez, R.R., Silvy, N.J., Frank, P.A. and Ransom, D., Jr. (2006), Comparison of Camera and Road Survey Estimates for White-Tailed Deer. The Journal of Wildlife Management, 70: 263-267. https://doi.org/10.2193/0022-541X(2006)70[263:COCARS]2.0.CO;2

[2] Sarmento, P., Cruz, J., Eira, C., & Fonseca, C. (2009). Evaluation of camera trapping for estimating red fox abundance. Journal of Wildlife Management, 73(7), 1207-1212. doi:10.2193/2008-288

[3] Kluever, B. M., Gese, E. M., Dempsey, S. J., & Knight, R. N. (2013). A comparison of methods for monitoring kit foxes at den sites. Wildlife Society Bulletin, 37(2), 439-443. doi:10.1002/wsb.261

[4] RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

[5] Salort, C., Green, T. (2020). Photo recognition for 4-poster tick management system®.

[6] Tabak, M. A., M. S. Norouzzadeh, D. W. Wolfson, S. J. Sweeney, K. C. VerCauteren, N. P. Snow, J. M. Halseth, P. A. D. Salvo, J. S. Lewis, M. D. White, B. Teton, J. C. Beasley, P. E. Schlichting, R. K. Boughton, B. Wight, E. S. Newkirk, J. S. Ivan, E. A. Odell, R. K. Brook, P. M. Lukacs, A. K. Moeller, E. G. Mandeville, J. Clune, and R. S. Miller. (2019). Machine learning to classify animal species in camera trap images: Applications in ecology. Methods in Ecology and Evolution 10(4): 585-590

[7] Tabak, M. A., M. S. Norouzzadeh, D. W. Wolfson, S. J. Sweeney, K. C. VerCauteren, N. P. Snow, J. M. Halseth, P. A. D. Salvo, J. S. Lewis, M. D. White, B. Teton, J. C. Beasley, P. E. Schlichting, R. K. Boughton, B. Wight, E. S. Newkirk, J. S. Ivan, E. A. Odell, R. K. Brook, P. M. Lukacs, A. K. Moeller, E. G. Mandeville, J. Clune, and R. S. Miller. 2018. mikeyEcology/MLWIC: Machine Learning for Wildlife Image Classification (MLWIC) (Version v0.1). DOI: 10.5281/zenodo.1445736.

[8] Learn to code - for free. (n.d.). Retrieved April 07, 2021, from https://www.codecademy.com/?g_network=g&g_device=c&g_adid=351522045838&g_keyword=codecademy&g_acctid=243-039-7011&g_adtype=search&g_adgroupid=70946090375&g_keywordid=kwd-41065460761&g_campaign=US_Brand_Exact&g_campaignid=1955172604&utm_id=t_kwd-41065460761%3Aag_70946090375%3Acp_1955172604%3An_g%3Ad_c&utm_term=codecademy&utm_campaign=US_Brand_Core_Exact_Net+New+%28Auto+Tagging%29&utm_source=google&utm_medium=paid-search&utm_content=351522045838&hsa_acc=2430397011&hsa_cam=1955172604&hsa_grp=70946090375&hsa_ad=351522045838&hsa_src=g&hsa_tgt=kwd-41065460761&hsa_kw=codecademy&hsa_mt=e&hsa_net=adwords&hsa_ver=3&gclid=CjwKCAjwjbCDBhAwEiwAiudBy4TuS2Cp8KKAqUk24SrKAYHCYIL5ii0W4025DiXuplHCfbD1Zan6ShoCdZoQAvD_BwE

[9] Anaconda Software Distribution. (2020). Anaconda Documentation. Anaconda Inc. Retrieved from https://docs.anaconda.com/

[10] Wildlife Insights Beta. (n.d.). Retrieved April 07, 2021, from https://www.wildlifeinsights.org/

[11] National Weather Service Corporate Image Web Team. “National Weather Service Climate.” National Weather Service, 24 Oct. 2005, w2.weather.gov/climate/index.php?wfo=okx.