WildLens: Machine Learning Enhanced Fauna Census through Trail Camera Imagery and Wildlife Photography

Brendan Bushbaker
CSE / Senior / 33% / Model Creation
bushbak2@msu.edu

Monika Kanphade
CSE / Senior / 33% / Researching
kanphade@msu.edu

Sriram Seelamneni
CSE / Senior / 33% / Paper Editing
seelamne@msu.edu

Abstract

Fauna identification plays a pivotal role in wildlife research and ecosystem monitoring. With the proliferation of trail camera images, traditional identification methods have become labor-intensive and cumbersome. To address this challenge, we present WildLens, a convolutional neural network (CNN) model tailored for wildlife recognition. Drawing inspiration from pioneering work by organizations like *Cvedia* [1] and *Wild Me* [5], WildLens offers a streamlined approach to expedite the detection and classification of wildlife species.

Our research leverages two comprehensive datasets: the *Animal*-$10N$ [6] dataset, comprising ten species with $50,000$ training and $5,000$ testing images, and the Animal Image dataset, featuring 90 species and $5,400$ images. The development of WildLens involves a structured, iterative approach: firstly, image standardization and organization; secondly, model training using test images; thirdly, rigorous model evaluation through cross-validation; and finally, fine-tuning and optimization.

The model achieved an accuracy of approximately 93% on training datasets and 60% on testing datasets. The finished code can be found at [8].

1 Introduction

1.1 Motivation

The importance of wildlife census has never been more critical than it is today. The rapid growth of the human population, urban expansion, habitat destruction, climate change, and poaching have had a detrimental impact on wildlife populations worldwide. Conservation efforts rely heavily on accurate and timely population data for various fauna. In response, we introduce WildLens, a machine learning model for fauna recognition, offering a non-invasive approach to passive and continuous wildlife census.

Traditional census methods are slow and resource-intensive, posing challenges for conservation agencies already facing budget constraints. These methods involve manual counting and sifting through thousands of trail camera images, often plagued by false positives triggered by non-fauna movements. Conservationists spend countless hours on this tedious task, impeding their overall effectiveness. WildLens, our proposed machine learning model, streamlines fauna detection and classification, significantly reducing the time required for accurate population assessments.

Furthermore, the global expansion of human populations has introduced invasive species to new ecosystems, where they often lack natural predators, leading to the displacement of native species and resource depletion. Early detection of invasive species is critical for their successful removal. WildLens, when integrated with trail camera systems, offers a cost-effective, passive, and autonomous defense system for native habitats. This technology ensures the preservation of our remaining natural habitats, safeguarding their health and biodiversity.

1.2 Ethical Considerations

Numerous ethical considerations will be prioritized in the development and usage of WildLens. First, no animals were harmed during the image collection process. The images are from professional wildlife photographers and trail cameras. Additionally, WildLens will be exclusive for academic and research purposes. With the threat of poachers and hunters misusing WildLens, only researchers and conservationists may use the software.

Also, WildLens will ensure accurate results. Researchers will make decisions and perform actions based on the results from WildLens. If these results are inaccurate, researchers may choose the wrong course of action and cause harm to the environment.

2 Related Works

The idea of capturing wildlife in its natural habitat has been made a reality in countless ways, spanning a variety of professions in recent years. However, when computer science and machine learning come into the picture, the scope of what can be accomplished upon the shutter of a camera changes for the better. Technological advances have been made on a global scale in almost every industry imaginable and in wildlife conservation especially. Many companies including *Wild Me* [5], *Cvedia* [1], and *Google* [2] have begun pioneering ways in which planet Earth can benefit from the rage of artificial intelligence today, and have been doing so successfully.

Following a project by *Cvedia* on the basis of animal detection reveals the impact of deep learning on creating a safer society. One of the resources they source information from include security surveillance cameras, in which they use animal detection software to highlight the presence of entities and classify them so as to reduce the possibility of a false alarm. Another source comes from wildlife cameras in capturing the habitats unseen by the human eye. The ability to identify certain species through machine learning has allowed for cameras to "smart track" animals by saving battery life and reducing data transfers. These are only a few examples of how *Cvedia* contributes to the safety of organisms on Earth and a handful of accomplishments made by their network [1].

Wild Me [5] takes on another approach in the same direction of wildlife conservation using their software to track animal populations and reduce the load on human counterparts in identifying species at risk of mass extinction. They eliminate the need to physically tag animals and promote the use of citizen scientists' contributions.

Finally, *Google* incorporates aspects from a multitude of data reserves to compile accurate readings of wildlife in many areas. Ranging from remote to well-known land, *Google* has provided a visual map of wildlife trends and insights suited to aid in the tracking and preservation of many species. Furthermore, their artificial intelligence model boasts the ability to filter blank images and classify up to 1,295 unique species around the world [2].

3 Methodology

The primary goal is to create a robust image recognition system capable of identifying animals in wildlife photographs accurately. This system can be invaluable for automating species identification, aiding wildlife researchers, and contributing to biodiversity conservation efforts.

3.1 Dataset Selection

Our approach to developing an image recognition model commenced with the careful selection of a suitable dataset for training. We chose the *Animal-*$10N$ [6] dataset due to its robust collection of $50,000$ training images and $5,000$ testing images. What set this dataset apart was its diversity in terms of image quality, encompassing a wide spectrum from hand-drawn illustrations to manipulated and noisy images animals to high-quality photographs of wildlife, free from noise. This diversity was expected to provide our model with a comprehensive understanding of the distinctive features defining each creature. Furthermore, the inclusion of pairs of similar animals, such as Hamsters and Guinea Pigs, within the dataset was intentional. We recognized the importance of addressing similarities between animals early in model development, as the model is intended for scaling to include a broader array of species in the future.

In exploring our dataset selection, it may be beneficial to the development of the model to discard some of the cartoon-ish or manipulated images in order to better classify animals in the wild. Classifying a drawing is not a concern to our purpose, thus further data scrubbing may be explored so that our model will better fit the purpose it is intended for. However, this may cause the model to lose robustness in classifying images of animals where the subject is not clear or noisy.

3.2 Model Development

Model development consisted of a multi-step approach using the prototype-testing method of model creation. Thus, several models were constructed, trained and evaluated. Below in Figure (1) is a road-map for the development process, which ultimately lead us to our final model creation, *WL18*.

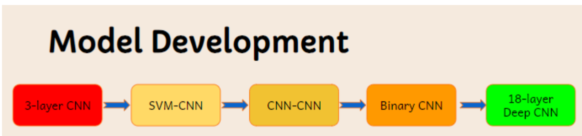

Figure 1: Prototype-Testing Development Process

3.2.1 8-Layer Model

Beginning the model development process, we initially considered a compact 8-layer Convolutional Neural Network (CNN). This rudimentary architecture comprised three convolutional layers and two interspersed max-pooling layers. At the end of the model, two dense, fully connected layers were implemented, culminating in a final dense layer with a softmax activation function. While serving as a foundational exploration, this model provided valuable insights that steered us towards more refined architectures for WildLens.

Despite its simplicity, the 8-layer model yielded suboptimal results, registering an accuracy below 50%. This underwhelming performance prompted a reassessment of our model's architecture, pushing us to explore the creation of more sophisticated structures for WildLens.

In response to the limitations of the initial 8-layer model, we started the development of more intricate architectures. These subsequent designs incorporated additional layers, diverse activation functions, and alternative configurations to enhance the model's capacity for feature extraction and classification. The iterative nature of the prototyping-testing process allowed us to systematically refine the WildLens model, aiming for superior accuracy.

3.2.2 Sifting Models

The next model in the progression of WildLens development was specifically tailored to sift through data categorized as "dirty." While the primary goal of WildLens was to classify real animals in natural settings like forests, fields, deserts, and jungles; the training data included drawings, tattoos, costumes, and other variant images that were irrelevant to the model's core purpose. An example of each classification is given in Figure (2). To address this issue, a dedicated model was conceived to discern between "real" and "fake" images, where "real" denoted authentic images featuring animals in expected settings, and "fake" encompassed variant images. This model would be used as the initial step in a multi-stage pipeline for training the final animal classification model, aiming to filter out irrelevant images and minimize bias.

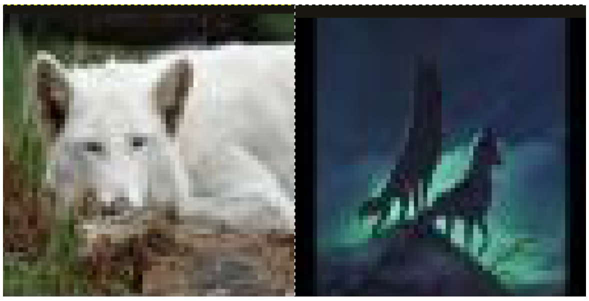

Figure 2: Example of a Real (left) and Fake (right) Image of a Wolf

The primary objective of this model was to perform binary classification, distinguishing between real and fake images. By learning this fundamental distinction, the model aimed to contribute to the creation of a more focused and precise dataset for subsequent stages of WildLens development. This strategic approach sought to eliminate sources of variance that could adversely impact the final model's accuracy and bias towards undesired results.

3.2.3 Labelling Software

The initial step in preparing the dataset for WildLens involved manually labelling real and fake data. To streamline this process, a GUI was developed, depicted in Figure (3). This GUI empowered researchers to classify a dataset comprising over $10,000$ images as either real or fake. The commitment to this labelling endeavor amounted to a cumulative effort exceeding 8 man-hours.

Figure 3: Custom Data Sifter

The custom GUI not only expedited the labelling process but also ensured a high degree of precision in categorizing images. Researchers could efficiently navigate through the dataset, making informed decisions about the authenticity of each image.

3.2.4 SVM Sifter

Following the completion of manual labelling, the development of the Support Vector Machine (SVM) architecture was initiated. Two fitting methods, Random Search and Grid Search, were employed to optimize the SVM model. The hyperparameter space was explored with tested values for margin softness (C) including $0.1, 1, 10, 100, 1000$, a range of gamma values $(1, 0.1, 0.01, 0.001, 0.0001)$, and multiple kernels such as Radial Basis Function, Polynomial, Sigmoid, and Linear. A total of 512 candidate combinations were assembled and ran through the search algorithms.

RGB images were initially formatted with dimensions of $64x64x3$. To align with SVM requirements, these images were flattened into column vectors of size $12,288$ before being fed into the SVM model.

Initially, it was anticipated that training 512 models would not impose an exorbitant time burden. However, as the training progressed and hours elapsed, it became evident that the available hardware lacked the required power for this task. Consequently, the SVM filter was discontinued. This decision reflects a practical response to hardware limitations, acknowledging the necessity of revisiting this approach in the future if enhanced computing power becomes available.

While the SVM Sifter was not viable with the current hardware constraints, it remains a prospect for exploration with proper computing resources. This choice reflects our prototyping approach to model development, where consideration is given to both dataset selection and model architecture.

3.2.5 CNN Sifter

To address the sifting classification task from a different perspective, the model described in Section 3.2.1 was repurposed for the binary classification task. While retaining the foundational architecture, certain modifications were introduced to tailor it for sifting. Notably, the kernel sizes of the convolutional layers were adjusted to $64, 128, 128, 256$, each accompanied by a Max Pooling layer and a Dropout layer set at 30% to mitigate overfitting. The final layer's activation function was switched from softmax to sigmoid, aligning with the binary nature of the output. For predictions, a threshold of 0.5 was established.

Initially, the model showcased promising performance, with an accuracy hovering around 81% and a low binary entropy cost of 0.46. However, a red flag emerged when the recall value was consistently at 1.00. This raised concerns about potential overfitting or convergence to a local minimum during the training process.

During training, this model was fused to the bottom of an eight layer model as described in 3.2.1. The CNN sifter played the role of gate-keeper into our primary model. We hoped that the filtered images would allow for better evaluation accuracy of the final model.

Further investigation, as discussed in the results section 4.2, revealed that the model was indeed fitting to a local minimum. The overemphasis on recall, while seemingly impressive, indicated a lack of generalization capability, ultimately rendering this model unsuitable for the project's objectives.

In light of these findings, the CNN Sifter model was regrettably dropped from the project. This decision was made in order to continue our objective from a different angle. The experience gained from this instance informed the selected model development strategy and reinforces the importance of thorough evaluation and interpretation of model metrics.

3.2.6 Binary and Categorical Fused Models

In an attempt to enhance accuracy through a systematic pipeline, the Binary and Categorical Fused Models represent an innovative approach. The methodology involves training separate categorical models on disjoint sets of animal classifications and subsequently utilizing binary classification models to further refine predictions. This fusion aims to harness the strengths of both categorical and binary models for improved accuracy in wildlife recognition.

Two categorical CNN models were trained on distinct sets of animal classifications, carefully chosen to include dissimilar animals within each set. The architecture of these CNN models closely resembles the description in Section 3.2.1. The selected sets were (Cat, Wolf, Cheetah, Chimpanzee, Hamster) and (Lynx, Coyote, Jaguar, Orangutan, Guinea Pig), ensuring that each set comprised animals with minimal visual similarity.

Four binary classification models, sharing a similar architecture with the categorical CNNs, were created. The final dense layer of each binary model was reduced to a size of one, with the activation function shifted to sigmoid. Each binary model was designed to classify pairs of similar animals, such as (Cat, Lynx) and (Wolf, Coyote).

Training and testing data were segregated accordingly, with each model being trained on data specific to the animals it was responsible for classifying. This meticulous organization aimed to ensure that each model specialized in recognizing the distinctive features of its assigned animal classes.

The prediction pipeline began with an image entering both categorical models. The predictions from these models were cross-examined, and the prediction with the highest confidence was chosen to proceed to the next phase. Based on the prediction made by the categorical models, the image was then fed into the appropriate binary classification model, ultimately yielding the final classification.

3.2.7 WildLens Layer 18 (*WL18*)

As the reader may gather from the title of this section, *WL18* represents the culminating model in the creation of WildLens. Drawing inspiration from the notable work presented in [7], which emphasized the effectiveness of deeper architectures in image classification, *WL18* aspired to achieve heightened performance through increased depth.

WL18 is an eighteen layer network whose architecture is depicted in figure (4).

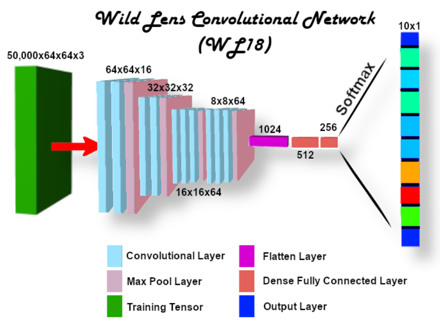

Figure 4: WildLens Layer-18 Architecture

Key Features and Innovations

- **Deeper Architecture**:
  WL18 is an eighteen-layer network, embracing the principle that deeper networks tend to excel in image classification tasks. The architecture is structured to capture intricate features and patterns within wildlife images.

- **Larger Kernel Sizes and Dense Layers**:
  This model diverges from its predecessors by employing larger kernel sizes and denser layers. The objective is to enhance the model's capacity for feature extraction and representation, facilitating more nuanced and accurate wildlife classification.

- **Three Filters and Padding in Convolutional Layers**:
  Each convolutional layer incorporates three filters and padding, contributing to a more comprehensive analysis of spatial features in the input images. This design choice aims to bolster the model's ability to discern intricate details crucial for wildlife identification.

- **Variable Learning Rate with ADAM Optimizer**:
  The model incorporates a variable learning rate strategy in conjunction with the ADAM optimizer. This adaptive learning rate is employed alongside validation loss plateau detection. If there is no decrease in the validation loss over the last 7 training epochs, the learning rate is adjusted, aiding the model in navigating towards a minima during training.

- **Model Saving Based on Validation Score**:
  Only the best-performing model, determined by the validation score, is saved during the training session. This meticulous approach ensures that the model's final state reflects optimal performance on the validation dataset.

4 Results and Analysis

The final results of this work were a testament to our perseverance to explore a new machine learning model and training methods while also continuously striving for a more accurate model. The final model had testing results with the following results presented in Table (1) below.

Table 1: Final Evaluation Results of WL18

Loss	Accuracy	F1-Score	Precision	Recall
1.201	59.5 %	56.7 %	71.4 %	47.3 %

4.1 8-Layer Model

The 8-Layer model, as detailed in Section 3.2.1, displayed promising capabilities for wildlife image classification using a Convolutional Neural Network (CNN). Initial accuracy results indicated approximately 48% accuracy when evaluating the model on testing datasets. The confusion matrix, depicted in Figure 5, provides a more granular understanding of the model's performance.

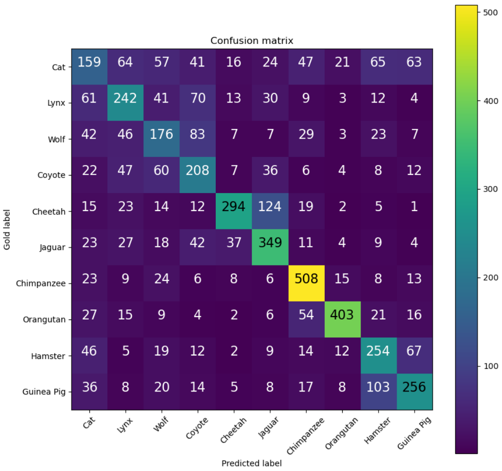

Figure 5: 8 Layer CNN Model's Confusion Matrix

This matrix showed that the trained model had powerful classification promise for Primates specifically, as evidenced by the high values along the diagonal corresponding to primates (Chimpanzee and Orangutan). This suggests that the model effectively learned distinctive features for this category. Canines (Wolf and Coyote) and Felines (Cat and Lynx) also exhibited strong classification ability. The high values along the diagonal for these categories indicate the model's proficiency in distinguishing between different species within these groups. The confusion matrix highlights challenges in accurately classifying household rodents (Hamster and Guinea Pig). The misclassifications and lower values in these categories suggest that the model struggled to discern subtle differences between these animals. Possible noise may have made this subset difficult to generalize.

4.2 CNN Filter Fused Model

The CNN Filter Fused Model, developed using the methods described in Section 3.2.5 and incorporating aspects from Section 3.2.4, faced challenges during the initial training and evaluation phases. Notably, the metrics of this fused model did not show improvement, and the model's performance was worse than the original 8-Layer model.

During investigation, a crucial insight emerged: a significant imbalance in the distribution of "real" and "fake" images in the manually labeled dataset. Specifically, 81% of the data was labeled as "real" images, while only 19% constituted "fake" images. This distribution imbalance rendered the model incapable of effectively generalizing between "real" and "fake" images.

The imbalanced dataset, with a sparse representation of "fake" images, posed a non-trivial challenge. The observed accuracy of 81% was, in fact, equivalent to random chance given the dataset's inherent bias toward "real" images.

Recognizing the limitations posed by data sparsity and time constraints, the decision was made to drop the filtering method from production. This choice reflected our team acknowledging the need to explore alternative strategies for handling the noisy dataset and improving model performance.

In summary, the CNN Filter Fused Model faced challenges stemming from an imbalanced dataset, highlighting the importance of addressing data sparsity and distribution issues to achieve more accurate and reliable results in wildlife image classification.

4.3 Systematic Pipeline (Categorical $\rightarrow$ Binary) CNNs

The systematic pipeline involving Categorical to Binary CNNs aimed to enhance accuracy through a carefully structured approach. The models exhibited promising trends during training, as illustrated in Figure (6), with both categorical models achieving metrics around 75%.

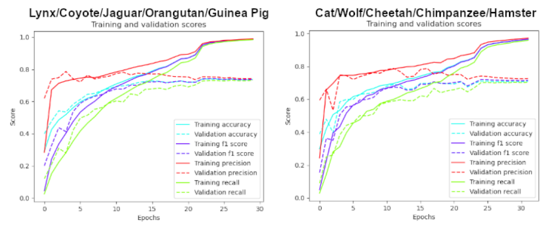

Figure 6: Training and Validation Trends Over Accuracy, F1-Score, Recall and Precision

The categorical models demonstrated a capacity to converge to a minima with higher degrees of confidence across multiple metrics. This distinction between different animal sets contributed to their successful performance.

Both models had metrics at or around 75%. This was promising as the binary models were trained next. These models seemed to show some difficulty in discerning between similar animals. As shown in the confusion matrices below, the binary classification task was a much more difficult one to achieve.

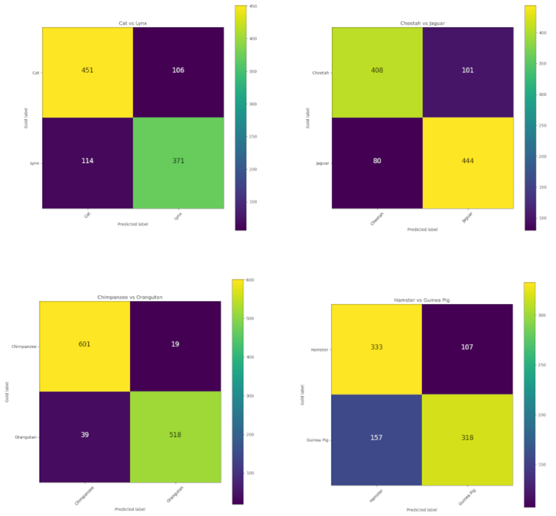

Figure 7: Binary Classification Confusion Matrices. In order from left to right, top to bottom: Lynx vs. Cat, Cheetah vs. Jaguar, Chimp vs. Orangutan, Hamster vs. Guinea Pig

The most difficult subset to train was the Hamster vs. Guinea Pig model. This was difficult due to the closeness in appearance of these two animals. The Chimp vs. Orangutans model enjoyed extremely satisfactory performance with a 95% accuracy rate.

The models were then fused by using custom made pipeline logic programmed in python. The metrics given in Table (2) were produced by this model.

Table 2: Fused Model Evaluation Results

Loss	Accuracy	F1-Score	Precision	Recall
1.201	54.3 %	52.6 %	54.5 %	53.5 %

While a good increase from the prior models, this still showed that there was room to improve. The confusion matrix shown below in Figure (8) showed that the model excelled at classifying Primates, but struggled especially with classifying Felines and Rodents. This is hypothesized to have been caused by the noise in the data. Notably for Cats, Hamsters and Guinea Pigs their owners were often pictured with the animal. This may have caused the model difficulties in generalizing. This high variance with these classifications is shown on the extreme corners of the confusion matrix below.

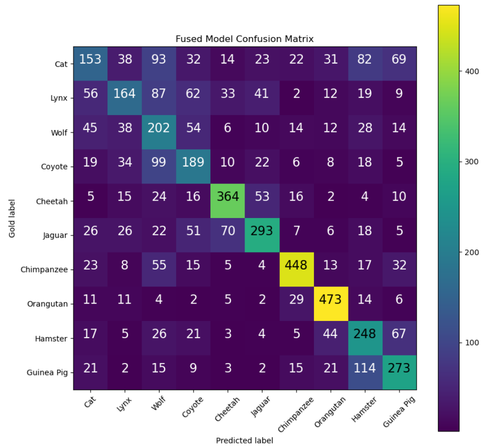

Figure 8: Fused Model Confusion Matrix

In the future, domestic animals may be dropped from the training dataset. There isn't a large need in the wildlife community to detect and census domestic animals.

This model overall had better performance, but there were other methods to test and better results to be sought.

4.4 WL18 Model

This 18-layer CNN model, as described in 3.2.7, is our final and best model yet. Below is the confusion matrix of this model,

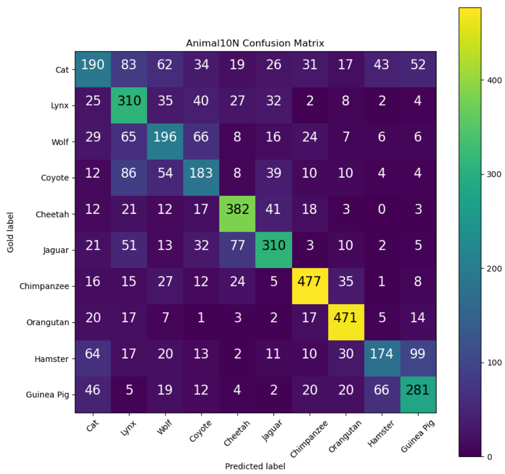

Figure 9: WL Model's Confusion Matrix

As you can see in the matrix, there is a significant increase in true positives for many species when compared to the 8-layer model. Although the model continued to struggle to classify between the rodents because of their similarities, it had got better at predicting the correct labels for other species.

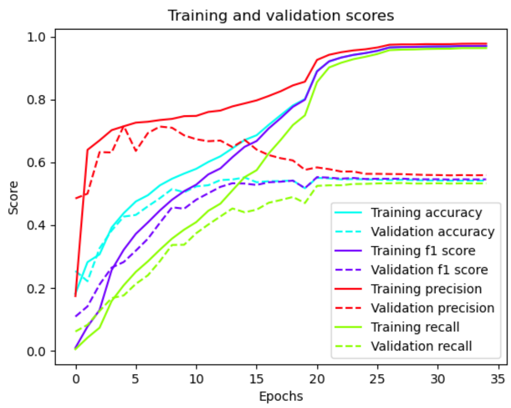

Figure 10: WL Model's Performance

The above graph illustrates the performance of our WL18 model over 35 epochs. It tracks six different metrics: training accuracy, validation accuracy, training precision, validation precision, training recall, and validation recall.

- Training and Validation accuracy: These curves represent how model correctly identifies positive and negative samples in the training and validation dataset respectively. Validation accuracy was around 59.5%, our best performance yet.
- Training and Validation precision: These curves indicate the proportion of positive identifications that were actually correct in training and validation datasets. Precision was around 71.4%, meaning when model predicts a positive result, it is correct 71.4% of the time.
- Training and Validation recall: These curves indicate the proportion of actual positives that were correct identified in the training and validation datasets. Recall was around 47.3%, meaning that of all actual positive cases , the model correctly identified 47.3% of the time.

5 Future work

The project has substantial potential with a range of unexplored possibilities and promising directions for future research. The following key points present potential areas of interest for future work:

- **Species Expansion:** The current model is designed to recognize and classify a specific set of species within our two datasets. Future work involves optimizing and expanding the model to accommodate a more extensive range of wildlife species. This expansion would involve collecting and annotating data for additional animal classes, thereby broadening the model's applicability to diverse ecosystems and contributing to its utility in real-world scenarios.
- **Continuous Model Training:** Implementing a continuous training approach will enable the model to adapt to evolving wildlife populations and environmental changes. Regularly updating the model with new data ensures that it remains relevant and accurate, especially in dynamic ecosystems where species compositions and behaviors may shift over time.
- **Ensemble Learning Techniques:** Exploring various learning techniques, where multiple models are combined to make predictions, offers a promising path for improving overall model performance. This approach can mitigate the impact of individual model biases and enhance accuracy, especially in scenarios where diverse data sources are available.
- **Ethical Considerations and Bias Mitigation:** As the model expands to include a broader range of species, addressing ethical considerations and potential biases becomes paramount. Future work involves implementing measures to mitigate biases in training data and ensuring the responsible use of the model in conservation and research efforts.

In summary, the future trajectory of our research involves not only optimizing the model for increased species recognition but also exploring innovative features and methodologies that contribute to the model's adaptability, accuracy, and ethical application in the realm of wildlife recognition and conservation.

6 Conclusion

On the journey to develop an animal recognition model, this project has witnessed good results, navigated challenges, and laid the groundwork for future explorations.We have created a Convolutional Neural Network (CNN) that is tailored for wildlife photography and a Data Sifter model that distinguishes real from manipulated images. As we reflect on this project, several key points include:

- **Model Development and Performance:** The CNN model, engineered for image recognition, has demonstrated very good performance on the *Animal*-10N [6]. Its ability to accurately classify ten distinct animal classes demonstrates the effectiveness of our architectural choices, optimization strategies, and the utilization of diverse training data.

- **Data Sifter Model:** The integration of the Data Sifter model, leveraging Support Vector Machines (SVM), addresses the challenges posed by non-relevant images. By sifting through and classifying images as real or fake, this model enhances the reliability and authenticity of our animal recognition system, especially in the context of wildlife photography.

- **Dataset Challenges and Solutions:** Our approach to dataset selection and preprocessing has been pivotal in overcoming challenges associated with noisy labels, irrelevant images, and diverse image types. The strategic curation of the *Animal*-10N [6] dataset, supplemented by additional data from Kaggle, ensures that our model is trained on a rich and varied set of wildlife images.

- **Future Directions:** The project's conclusion paves a path future exploration. The identified future work directions encompass a broad spectrum, from enhancing model optimization and scaling to diverse domains to exploring dynamic learning for evolving environments. These directions underscore the project's potential for growth and adaptation to emerging challenges.

- **Collaboration with Local Wildlife Associations:** A big problem during training of the models was noisy data. In the future, obtaining trail-cam photographs and videos from an official organization would improve the quality and target of our datasets. Future work would involve collaboration and data consumption from organizations with pre-labelled data of various faunas.

- **Contributions to Wildlife Conservation:** Beyond the model development, our project aspires to contribute meaningfully to wildlife conservation. This alignment with conservation initiatives opens avenues for deploying the model in biodiversity monitoring, supporting ecological research, and aiding conservation efforts on a global scale.

As we conclude, it is important to acknowledge that the field of image recognition in wildlife photography is dynamic and continually evolving. Our contributions, both in terms of model development and the thoughtful consideration of dataset challenges, lay a foundation to build upon. We are always looking for ways to improve our work and use it for the good of nature and wildlife. We are excited to see what the future holds for us and how we can make a difference in the world with our models. We are grateful for this opportunity and we hope to continue to learn and grow with technology and conservation in mind.

References

[1] Animal Detection | CVEDIA | AI Video Analytics for any hardware. (n.d.). https://www.cvedia.com/animal-detection

[2] Google (Sept. 27 2023) Home, Wildlife Insights. Available at: https://www.wildlifeinsights.org/ (Accessed: 27 September 2023).

[3] Ng, X. L., Ong, K. E., Zheng, Q., Ni, Y., Yeo, S. Y., and Liu, J. (2022). Animal Kingdom: A Large and Diverse Dataset for Animal Behavior Understanding (arXiv:2204.08129). arXiv. http://arxiv.org/abs/2204.08129

[4] Wildlife Insights. (n.d.). https://www.wildlifeinsights.org/

[5] Wild me. (n.d.). Wild Me. https://www.wildme.org/

[6] Animal-10N Dataset | Song, H., Kim, M., and Lee, J., "SELFIE: Refurbishing Unclean Samples for Robust Deep Learning," In Proc. 36th Int'l Conf. on Machine Learning (ICML), Long Beach, California, June 2019

[7] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. https://doi.org/10.48550/ARXIV.1409.1556

[8] https://gitlab.msu.edu/bushbak2/wildlens