International Journal of Scientific Research and Engineering Development-– Volume 5 Issue 4, July-August 2022

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Wildlife Classification On Camera Trap Images Using Deep Learning Models

Supritha R*, Rochana G**

**(Data Science, NMKRV college, Bangalore*

Email: supritha,gowda621@gmail.com)

*** (Data Science, NMKRV college, Bangalore*

Email: rochanagiridhar07@gmail.com)

--*************************----------------------------------

Abstract:

In recent years, automated surveillance systems called camera traps have helped conservationiststo monitor and study a wide range of ecology with less human interference. Camera traps aretriggered by motion or warm objects, and passively record the behavior of species in the areawithout disturbing their natural tendencies.However, camera traps also generate a huge amount of data which exceeds the capacity ofhumans to filter through. That's where machine learning and deep learning can help. Advances incomputer vision can help automate tasks like species detection, classification, and individualidentification so humans can more effectively learn and protect the ecology.

In this project, we will try to classify the objects for wildlife species. Classifying wildlife speciesis an important step to sort through images, quantify observations, and quickly find those withdifferent species.We'll use a pretrained EfficientNet model and ResNet model with more layers like ResNet50 toclassify the wildlife species and these models have been compared to find out which model givesus the best results. In the end EfficientNet model gave best accuracy than ResNet50.

Keywords: Wildlife image classification, pretrained models, ResNet50, EfficientNet.

--*************************----------------------------------

I. INTRODUCTION

In this paper, we address the task of wildlife classification from camera trap images using different pretrained models. Camera traps are aninstrument used by conservationists to learn and observe a wide range of ecologies with less human interference. However, they also generate a huge amount of data that quickly exceeds the capacity of humans to sift through. In such cases we can use different algorithms which will help automate tasks like species detection and identification of different species. There are 7 different species of animals in our dataset including blank images. We use different pretrained models to detect the species and compare the results to find which model gives us the most accurate results.

II. MOTIVATION

Our main aim is to enable a computer to solve tasks using machine learning and deep learning techniques. As per sources a larger amount of data is available on wildlife activity over time and space domain. Camera traps provides images of animals from their natural habitat. Using CNN, we can extract valuable information from the images to perform operations on the data.

III. OBJECTIVE

The model is built on dataset to predict the accurate and effective animal species with highest accuracy.

Using the accurate technique to predict the wildlife species with camera trap images which will help ecologists to find out the different species in in different location.

IV. BACKGROUND STUDY

1.In this paper, a deep learning-based model is used for automatic identification, description of wild animals in camera-trap photos. It will also count the number of wild animals. They trained the DNN to detect, sum up the number of animals, and give details of the behaviour of the animals using the Snapshot Serengeti dataset, which has 3.2 million photos of 48 different species. They added SoftMax output layer for each additional attribute to predict the probability of behaviour in the image. They trained neural network models which are AlexNet, VGG, GoogLeNet, and Resnet with different numbers of layers and compared them based on accuracy. The VGG model had the best accuracy of the nine neural networks, with a score of 96.85%. Their approach eliminated 98.4 percent of manual effort in animal identification.

2. In this paper, they have used camera trap images dataset which contains 22 different species of wildlife samples which are more than 1,00,000 in number. Considering the complexity of the landscape, different lighting conditions, and many other factors by using trail cameras, they improved the ability of systems to generalize new environments. The dataset also has many issues which include data imbalance, unlabelled data, and many other which helped them explore even in semi-supervised area. They focused on detecting endangered Florida panther.

A. *CAMERA TRAP IMAGES*

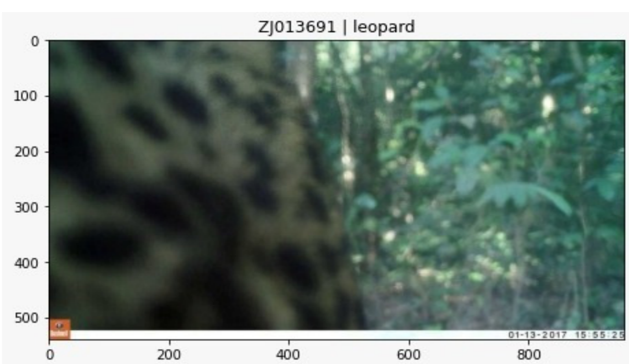
Camera trap images are images which are recorded only when there is any live moment of the objects. These camera traps are kept in the forest i.e., in the branch do that it records the moving objects immediately if it is detected. There can be even blank images captured sometimes due to the movement of leaves due to wind. These camera trap images help to find out the different species which in turn will be helpful in many studies. We can even find out the extinct animals in that locationand the number of species which is more in number.

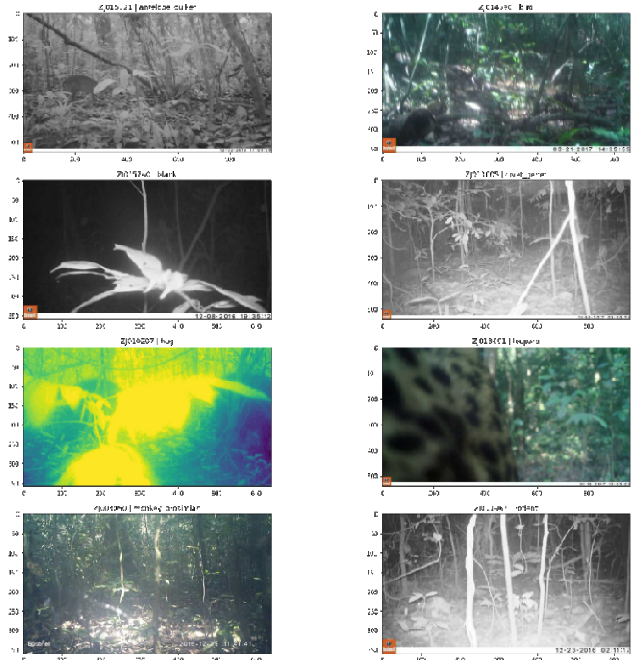
Species	Count
Monkey prosimian	2492
Antelope duiker	2474
Civet genet	2454
Leopard	2423
Rodent	2013
Bird	1641
Hog	978
blank	2213

V. *EXPERIMENTS*

a. *DATASET*

The dataset used in our project contains training and testing images where there is around 16,488 training images and 4,465 testing images. After training the model it is evaluated using test features, train features and train labels files which are in csv format.





VI. MODEL TRAINING

Before training the model, the photographs are split into train and eval sets. 25% of the information is kept for evaluation and stratify by the target labels to make sure we've similar relative frequencies of every class within the train and eval sets. Then the validation is finished to confirm that our split has resulted in roughly similar relative distributions of species across the train and eval sets. Images Dataset class is formed to define way to access the information and any transformations we would want to use.Then a collection of transformation is additionally defined using the special methods. These are applied to every image before returning it. These transformations transform the pictures into size 224×224 because the model ResNet50 which we're using is trained on images of size 224×224. So, we resize to same dimensions as ResNet model.
The dataset is loaded into a dataloader. The DataLoader class will iterate through the dataset in batches.

We'll use a pretrained EfficientNet and ResNet50 models as our backbone. ResNets are one ofthe most used networks for image classification tasks. The pretrained model outputs a 2048-dimension embedding, which we are going to connect to two more dense layers, with a ReLU and Dropout step in between.

The final layers defined are the new head of our model andpermit us to reform the image embeddings produced by the pretrained "backbone" into the 8-dimensional output required to find out the species classification task. We will instead prepare the model for the current task by redefining model to produce an 8-dimensional output corresponding to our 8 species classes (including blanks). And more layers are added in between. The ReLU layer introduces non-linearity into the model, which will activate important features and noise. And also the Dropout layer could be a commonly used regularization component that randomly drops some nodes from the previous layer's outputs (10% of nodes during this case).

We're using loss function i.e., Cross entropy loss or log loss which is employed for multi-class image classification. This loss function is employed to compute the loss for each training batch and to update the parameters accordingly. The loss gradually goes down with the amount of epochs.

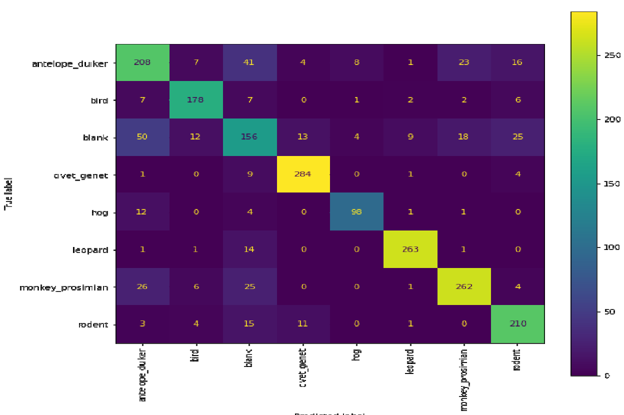
VII. PROPOSED ARCHITECTURE

We train a baseline model using ResNet-50 architecture. The dataset comprising of 16,488 training images and 4,465 images testing. We have also employed cropping, flip, and rotation for data augmentation. Images resized to 224 × 224px are processed through the model in a batch size of 34, using a learning rate of 0.001. We achieve a baseline test accuracy of 80.75% on our ResNet-based model.

And the second model is EfficientNet architecture. Then the network is fine-tuned for obtaining maximum accuracy but is also penalized if the network is very computationally heavy. It is fine for slow inference time when the network takes a lot of time to make predictions. The architecture uses a mobile inverted bottleneck convolution like MobileNet V2 but is much larger due to the increase in flops. EfficientNet model gave accuracy of 95.63% and works better than ResNet-50.

VIII. EVALUATION

To test the implementation and to obtain the results of the objectives, many experiments are conducted. ResNet-50, EfficientNet are the architectures on which the evaluation results are based. F1-score, Precision, Recall, Accuracy and Confusion Matrix are the key features in evaluation. The metrics are used to classify animal species.



IX. CONCLUSION

Detection of animals from the dataset would be of great importance to study the detail of animal species. Hence, evolving technology of various DL algorithms can be put into animal identification and detection. We used pretrained models such as ResNet-50 and EfficientNet model. These models' results were compared to check which gives better results with higher accuracy. After the results were obtained, it is concluded that EfficientNet was the best for animal classification with an accuracy of 95.63% where ResNet-50 gave 80.75%. In the end we conclude that EfficientNet was better pretrained model compared to ResNet-50 with higher accuracy.

X. REFERENCES

[1] Benjamin Kellenberger, Michele Volpi, Devis Tuia,. Fast animal detection in uav images using convolutional neural networks.

[2] Tanishka Bafhe, Janhavi Borde, Bhagyashree Waghmare, Vaishnavi Thakur, Anagha Chaudhari. Stusy of deep learning algorithms to identify and detect endangered species of animals.

[3] Identifiaction and classification of wildlife from camera-trap images using machine learning and computer vision.

[4] Crystal Gagne, Jyoti Kini, Daniel Smith, Mubabrak Shah. Florida wildlife camera trap dataset.

[5] Agnieszka Miguel, Sara Beery, Erica Flores, Loren Klemesrud, and Rana Bayrakcismith. Finding areas of motion in camera trap images. In 2016 IEEE international conference on image processing (ICIP), pages 1334–1338. IEEE, 2016.

[6] Mohammad Sadegh Norouzzadeh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S Palmer, Craig Packer, and Jeff Clune. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proceedings of the National Academy of Sciences, 115(25):E5716–E5725, 2018.

[7] Alexander Gomez Villa, Augusto Salazar, and Francisco Vargas. Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. Ecological informatics, 41:24–32, 2017.

[8] Matthew Kutugata, Jeremy Baumgardt, John A. Goolsby.Automatic camera-trap classification using wildlife-specific deep learning in nilgai management.

[9] Shreyas Sreedhar, Sandesh S, Prarthana P, N Karthik Pranav, Prabhanjan S, Srinidhi K. A literature survey on wildlife camera trap images processing using machine learning techniques.

[10] Christopher Beirne, A. Cole Burton. Use of object detection in camera trap image identification: assessing a method to rapidly and accurately classify human and animal detections for research and application in recreation ecology.

[11] N. Banupriya, S. Saraya, Rashi Swaminathan, Sachinthaa Harikumar, Sukhita Palanisamya, Animal Detection using Deep Learning Algorithm, Journal of Critical Reviews, 2019.