Dynamic Detection of Wild Animals Utilizing FGVC8 Images

Kompella Bhargava Kiran(✉), Veera Bhargavi Tadala, Naga Varshitha Yaragani, Anjali Puranam, and Sri Sai Pavan Velugubantla

Vishnu Institute of Technology, Bhimavaram, Andhra Pradesh, India
bhargav.kiran@gmail.com

Abstract. Using YOLOv5 for object detection and annotation, the research focuses on counting individual animals in camera trap sequences. This is summarized in the abstract. To increase the precision of animal counting models, the research integrates data from several sources, including Landsat-8 images, I Naturalist, and camera trap data. Data augmentation and fine-tuning procedures are used to solve challenges in wildlife monitoring, such as lighting fluctuations and sparse temporal samples. YOLOv5 is used to provide annotations that are crucial for training animal counting models and for precise item detection and compared with YOLOv8 model to predict which model got the best result and observed YOLOv8 got more accuracy. The project advances wildlife monitoring by improving models, incorporating new data, and exploring cutting-edge methods for accuracy and scalability. Insights that greatly support conservation and animal monitoring initiatives. Technological developments, data fusion, assessment criteria, deep learning applications, and historical viewpoints from prior contests are important points. In increasing the use of cutting-edge deep learning techniques in wildlife monitoring and conservation initiatives. For object detection, and image classification, participants can experiment with cutting-edge technologies. This emphasis on utilizing cutting-edge methods improves the precision and effectiveness of biodiversity assessment procedures. The project's findings, which include tracking wildlife populations, analyzing behavior, and keeping an eye on habitat changes, can support ecological research. Main objective is making wise decisions on conservation efforts requires knowledge of this information. The Realtime example use cases where we can use iwild cam is used are biodiversity monitoring, habitat monitoring.

Keywords: Wild animal detection · YOLOv8 · YOLOv5 · Deep learning · Computer Vision

1 Introduction

The state of nature is dire, and it won't get any better. More than a million species are in danger of going extinct as a result of a rate of extinction not witnessed in ten million years [1]. The global issue of animal extinction must be addressed before it gets out of

P. Pareek et al. (Eds.): IC4S 2024, LNICST 599, pp. 124–133, 2025.
https://doi.org/10.1007/978-3-031-77081-4_11

control. Experts predict that the current rate of rapid species loss is 1,000–10,000 times higher than the natural extinction rate, according to The World-Wide Fund for Nature [2]. It is vital for biodiversity to stop extinction. The extinction of any species can upset the ecosystem's delicate equilibrium because they are all important to it [3]. Our goal is to create a system that uses camera trap photos to identify and categorize different species of wild animals. This supports conservation efforts by assisting researchers in understanding wildlife populations, behavior, and habitat use. The main purpose of this project is to monitor the camera trap images taken from the fgvc8 camera images data [4]. Data from the camera traps can offer important insights about the number, distribution, and behavior of animals as well as the potential and risks associated with conservation. Camera trap data do, however, also provide a number of difficulties, including vast data sets, a variety of uncommon species, inconsistent image quality, and intricate analytical techniques. By integrating multispectral images, cloud computing, artificial intelligence, and standardized data, this work seeks to overcome some of these issues and increase the precision and effectiveness of camera trap data analysis [5]. By uploading the data to GBIF, a global platform for biodiversity data, the project also aims to promote cooperation and data exchange among camera trap researchers and practitioners5. By doing this, the project seeks to improve the management and analysis of camera trap data, further our understanding of animals, and promote wildlife conservation. Camera trap data analysis still has a lot of unanswered questions and difficult problems that call for more investigation [6].

How to handle the long-tail distribution of data from camera traps, in which certain species are extremely rare and others are extremely frequent. This may have an impact on population estimate accuracy as well as item detection and classification model performance.

How to enhance the tracking and identification of individual animals both within and between image sequences, with a focus on species that lack distinguishing traits or marks. Estimating the density, abundance, and survival rates of wildlife populations depend on this. The integration of environmental and contextual data, including meteorological information, multispectral imaging, habitat attributes, and human activities, into the analysis of camera trap data. Understanding the factors influencing animal distribution and behavior, as well as the effects of climate change and human disturbance, can be aided by this number [7]. How to guarantee camera trap data and metadata quality, standardization, and interoperability; additionally, how to handle the moral and legal implications of data sharing and publication. This can help with data reuse for different uses, meta-analysis, and cross-project collaboration. It is important to complete this project because by doing this research we can monitor the animal insights and we can also determine the count of each species in that location. If we get the count of each species, this data can be useful to ecologists so, that they can monitor the biodiversity and can make an awareness about animals going extinct. We propose the solution to count and detect the animal species in the given camera trap images.

2 Literature Survey

The course "dynamic detection of wild animals using deep learning" addresses several different subjects. These include of previous iWildcam competitions, species detection, counting, deep learning models, and assessment tools. Beyond conventional computer vision methods, the course investigates deep learning strategies including convolutional neural networks (CNNs) [8–10] and multimodal data fusion. It also explores the use of deep learning in biodiversity monitoring and evaluation indicators such as Mean Column wise Root Mean Squared Error (MCRMSE). The literature review integrates insightful observations from prior iWildcam competitions, specifically iWildcam 2018, 2019, and 2020. These insights provide insightful details about the process, difficulties encountered, and development of the course "dynamic detection of wild animals using deep learning" addresses several different subjects. Solutions gradually.

The literature evaluation for the iWildcam 2021 - FGVC8 project examines a broad variety of studies in numerous disciplines that are essential to meeting the competition's goals. An important field of research is the processing of camera trap data, where scientists look into methods to deal with issues such as sparse temporal samples and motion-triggered bursts. The paper also looks at methods for accurate species identification, population density estimation, and effective data processing. The literature investigates many techniques for species recognition and counting, such as convolutional neural networks (CNNs) and conventional computer vision techniques. Through multimodal data fusion, the integration of camera trap data with other sources, like multispectral imaging and ecological metadata, is also investigated. The evaluation metrics utilized in research projects and competitions for biodiversity monitoring are also reviewed, with a particular focus on the decision to adopt Mean Column wise Root Mean Squared Error (MCRMSE) as the assessment statistic for iWildcam 202. Finally, the use of deep learning to biodiversity monitoring is the main topic of the literature review.

It is employed to investigate different research fields that are necessary to meet the competition's goals. It is centered on camera trap data analysis, tackling issues such as sparse temporal samples and motion-triggered bursts. Effective data processing, precise species identification, and population density calculation techniques are also reviewed. The literature review for the iWildcam 2021 - FGVC8 project covers a broad range of topics that are essential to accomplishing the competition's goals. The processing of camera trap data is a crucial component in which scientists concentrate on problems such as sparse temporal samples and motion-triggered bursts [11]. Effective data processing, precise species identification, and population density estimation techniques are also examined in this research. The literature study looks at several methods for identifying species. It thoroughly investigates traditional computer vision methods as well as the advancements in deep learning techniques, specifically convolutional neural networks (CNNs).

3 Design and Analysis

Similar works that exist in same field have been solved but have some shortcomings. Through meticulous analysis, we were able to identify the benefits and limitations of several of the current works. The YoloV3 model and video data are used in "Wild animal detection and identification using deep learning" [12, 13] to identify and label the names and count of wild animals. However, there are some drawbacks, including the possibility that the count may not be accurate because the animals roam around and that the camera may catch the same animal more than once, which could lead to an increase in the count. Thus, it might not be precise. Let's move on to "Wildlife surveillance using deep learning methods" [14]. Upon identifying an animal, the surveillance camera takes a picture and even applies two deep learning techniques. The only problem is that it doesn't label the animal's name or provide the precise location of the animal in the picture. In contrast, the project "Animal Detection using Deep Learning Algorithm" [15] uses video footage to identify animals when they are spotted on roadways in order to prevent collisions between them and cars. When an animal is seen in the video for this project, a picture of that specific animal from the dataset is displayed. The project's drawbacks include the possibility of human and animal injury if someone is unaware of the type of animal present. Additionally, the feature "Wild animal intrusion detection using deep learning techniques" [16] detects the presence of wild animals and releases loud noises that startle them into leaving, though not all animals will be alarmed by these loud noises. Some might become aggressive and hurt others.

Numerous techniques have been employed to identify animals for a variety of goals, however each method is either lacking something or in a certain area. They are all neither flawless nor flawed. All they need to improve in certain domains. While comprehending these previous works, several inquiries might come up like,

1. Which techniques are used?
2. What are the advantages of these work?
3. What kinds of technological tools are available to improve accuracy?
4. How can the project be applied to real-world experiences?

When analyzing these previous efforts, the majority of them made use of video data, which may be utilized to capture any kind of wild animal, however the results might not be entirely accurate. Thus, they must employ several techniques to increase their accuracy.

1. What is the reason for using video data?
2. What efforts should be made to increase accuracy?
3. Are these works useful in protecting people or animals from harm?
4. Are there any changes to be made?

In our investigation, we identified wild animals in the provided photographs and labeled them with their names to determine what type of animal they were. We used the YOLOv5 [17, 18] model and image data for detection because it contains an inbuilt Blackbox that lets us find an animal in the image by providing the requested place and name in the image. We also used image data for detection. Figure 1 shows the brief design of our process.

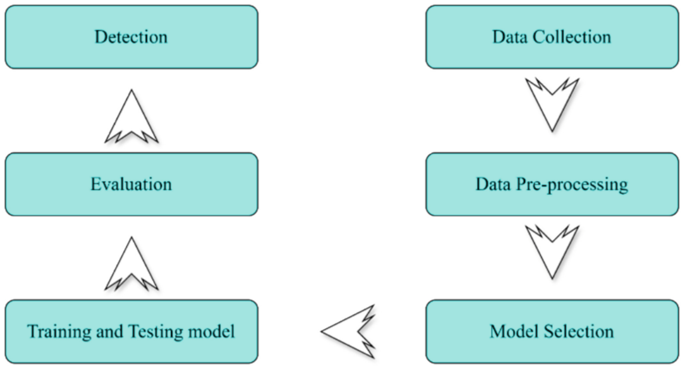

Fig. 1. Flowchart representing methodology

4 Experiments and Results

For implementing this project, we have collected the dataset from the Kaggle competition named iWildCam 2021 - FGVC8, the size of the dataset is 112 GB and it includes a total of 419709 files. Annotated photos of several wildlife species captured with camera traps can be found in the Kaggle competition dataset. It is quite large, has a wide range of species represented, and might have gone through preprocessing operations like normalization and resizing. The data is given in the format of JPG PNG files and JSON files format. For each image, a JSON file is given according to that which represents the annotation which is related to the location. By using the JSON format files we have trained our model to get annotations. For this project, we have used the YOLOv5 model to train and get the results. Coming to the procedure we have developed this project using the Google Colab notebook for implementing the code. First, we have uploaded the dataset into the drive hence we can use it in the colab. First, we have cloned the YOLOv5 repository into our collab notebook to use the YOLOv5 model. Next for setting up the YOLOv5 environment, we have installed the necessary dependencies. We have found that there are total 382 images with 198 classes having 593 instances.

We have imported several modules like torch python.display to display the images and utils.downloads to download the models and datasets. After installing the necessary libraries and modules we have mounted the google drive for the specific path. Next, we have displayed the YAML content. The next step is to define the model configuration and the architecture in this we have extracted the total number of classes of animals provided. We have retrieved the necessary parameters to train and evaluate our model. We have enabled to write the content of a cell potentially containing variables in our specified file. The next step in our procedure is to fire off the training. We passed a different number of arguments such as img to define the input image, batch to determine the batch size, epochs to define the number of epochs usually they are 3000 plus are common here, data to set the path to yaml file, cfg to specify our model configuration, weights to specify the custom path to weights, name variable represents the resulting naming, we have also used nosave parameter which saves only the final checkpoint and cache for faster image training. In our implementation, we have enclosed with hundred epochs.

After training the dataset with the given JSON formatted files, we get the annotations around the animal that is detected in the given image. In this competition, they are also given the test dataset which contains the images of the fgvc8 camera. By using this dataset, we have performed training and calculated the evaluation metrics. When evaluating the precision, recall, and F1-score of wildlife detection metrics, they are critical. They aid in evaluating the models' classification accuracy, which is essential for comprehending how well the animals perform in conservation-related tasks. As in [19] we also used different evaluation measures and determined by plotting the different graphs. From these graphs, we have calculated the accuracy. We have used measures such as box loss, obj loss, cls loss, precision, and recall. We have also used the epoch data to calculate the accuracy of our model. After performing the evaluation, we got the following graph results which shown in Fig. 2.

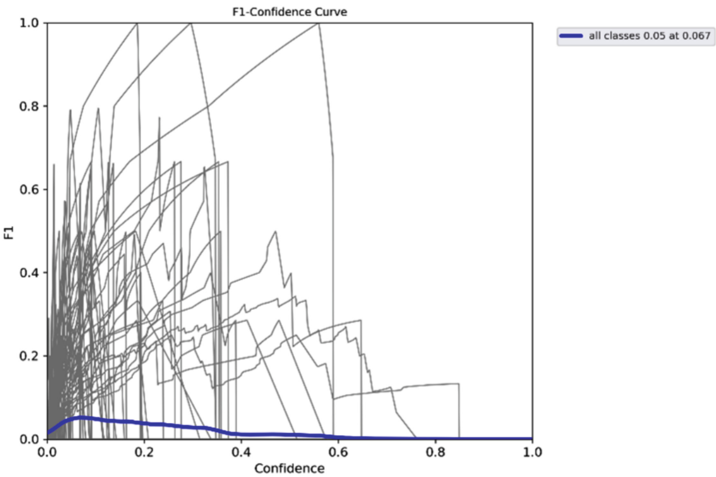

Fig. 2. Confidence curve of YOLOv5

This is the resultant F1 score curve [20] which is plotted against confidence for all classes we got 0.05 F1 score at 0.067 confidence. That is a good score with respect to the confidence parameter. F1 score nothing but an evaluation metric that measures models accuracy, good F1 score represents good accuracy of model [21].

After getting the results of YOLOv5 model we also trained the model using the latest version of yolo that is YOLOv8. We got the results as shown in Fig. 3. In this implementation first we have imported the necessary modules. After completing the setup of YOLOv8 [22] we started the custom training of images. It is similar to custom training of YOLOv5. We have taken 25 epochs to determine the accuracy. Then we got the result of F1-confidence curve as follows. In that we observed that for all classes we got 0.181 F1 score at 0.10 confidence.

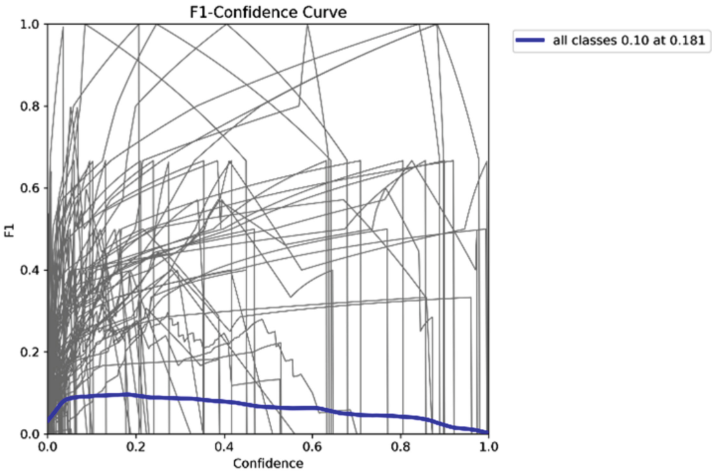

Fig. 3. Confidence curve of YOLOv8

After getting the epochs result, we compared the both models accuracies respective the map, map value is nothing but it is a catalogue object that provides mapping between time and epochs.

Table 1. Mean average precision, precision, recall and F1score of YOLOv5 and YOLOv8

Model	map (50)	map (50–95)	precision	recall	F1score
YOLOv5	0.316	0.152	0.651	0.268	0.378
YOLOv8	0.16	0.0982	0.694	0.125	0.449

From Table 1 results we observed that for YOLOv5 we got precision of 0.651, recall of 0.268 by calculating the F1 score we got 0.378 for 200 epochs at map@50 of 0.316 and for YOLOv8 [18] model we got precision of 0.694 recall of 0.125 by calculating F1 score we got 0.449 for 25 epochs at map@50 of 0.16. We can analyse that YOLOv8 model got better precision for just 25 epochs but at the same time we have to run 200 epochs for getting similar precision in YOLOv5 model. YOLOv8 is quicker and more precise than YOLOv5, yet YOLOv5 is simpler to use. YOLOv8 is a superior option, nonetheless, for applications that need real-time object detection.

We got the detection of animals after applying the model as shown in Fig. 4.

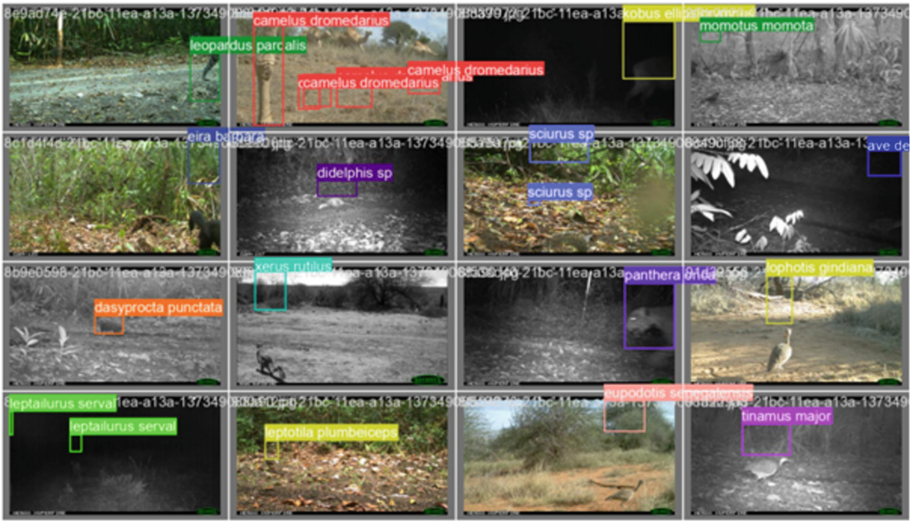

Fig. 4. Sample detection of wild animals using YOLOv5 and YOLOv8

5 Conclusion

Taken the vast dataset from the Kaggle competition and detected the each and every species which is presented in the image. The main purpose of this project is to detect the wild animals given in the camera trap images, in addition to that we also classify the different species presented in that particular image. In this detection process we have compared the results of two models YOLOv5 and YOLOv8 and compared the evaluation metrics such as map@50, map@(50–95), precision, recall and F1 score. By comparing these two models we observed that YOLOv8 model works better than YOLOv5 model for this dataset.

The results of the iWildCam 2021 - FGVC8 project could impact future directions, approaches, and technological advancements in the field of computer vision for ecological investigations. Modern machine learning techniques must be applied to complicated ecological concerns, and the challenges this competition presents are evidence of the interdisciplinary collaboration needed. The accuracy and effectiveness of wildlife detection systems can be improved by incorporating cutting-edge computer vision and machine learning techniques, which will further the field's continuing study.

References

1. Stork, N.E.: Re-assessing current extinction rates. Biodivers Conserv. **19**, 357–371 (2010)
2. Anyango-van Zwieten, N., Lamers, M., van der Duim, R.: Funding for nature conservation: a study of public finance networks at World Wide Fund for nature(WWf). Biodivers. Conserv. **28**, 3749–3766 (2019)
3. Fischman, R.L., Meretsky, V.J., Castelli, M.P.: Collaborative governance under the Endangered species act: an empirical analysis of protective regulations. Yale J. on Reg. **38**, 976 (2021)

4. Beery, S., Agarwal, A., Cole, E., Birodkar, V.: The iwildcam 2021 competition dataset. arXiv preprint arXiv:2105.03494. (2021)
5. Bühne, H., Pettorelli, N.: Better together: integrating and fusing multispectral and radar satellite imagery to inform biodiversity monitoring, ecological research and conservation science. Methods Ecol. Evol. **9**, 849–865 (2018)
6. Villa, A.G., Salazar, A., Vargas, F.: Towards automatic wild animal monitoring: identification of animal species in camera-trap images using very deep convolutional neural networks. Ecol. Inform. **41**, 24–32 (2017)
7. Pimm, S.L., et al.: The biodiversity of species and their rates of extinction, distribution, and protection. Science (1979), **344**, 1246752 (2014)
8. Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J.: A survey of convolutional neural networks: analysis, applications, and prospects. IEEE Trans. Neural Netw. Learn. Syst. **33**, 6999–7019 (2021)
9. Bonthu, S., Kiran, K.B., Deenakonda, M., Rao, V.V.R.M., Jagadeesh, S.: Multi-label and multi-class classification on a custom dataset using convolution neural networks. In: 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 576–579 (2023)
10. Ch SNLS Sai Baba, B., Swathi, M., Kiran, K.B., Bharathi, B.R., Matta, V.D., Veenadhari, C.H.L.: Medical plants identification using leaves based on convolutional neural networks. In: International Conference on Cognitive Computing and Cyber Physical Systems, pp. 163–171 (2023)
11. Findlay, M.A., Briers, R.A., White, P.J.C.: Component processes of detection probability in camera-trap studies: understanding the occurrence of false-negatives. Mamm Res. **65**, 167–180 (2020)
12. Gabriel, M., Cha, S., Al-Nakash, N.Y.B., Yun, D.: Wildlife detection and recognition in digital images using YOLOv3. In: 2020 IEEE Cloud Summit, pp. 170–171 (2020)
13. Balaji, R., Prabaharan, G., Singh, A.R., Athisayamani, S., Sarveshwaran, V., Daniya, S.: Multi-scale features fusion with YOLOv3 for detecting small and fine tumors in MRI images. In: 2022 6th International Conference on Electronics, Communication and Aerospace Technology, pp. 1545–1549 (2022)
14. Chen, R., Little, R., Mihaylova, L., Delahay, R., Cox, R.: Wildlife surveillance using deep learning methods. Ecol. Evol. **9**, 9453–9466 (2019)
15. Banupriya, N., Saranya, S., Swaminathan, R., Harikumar, S., Palanisamy, S.: Animal detection using deep learning algorithm. J. Crit. Rev. **7**, 434–439 (2020)
16. Sabeenian, R., Deivanai, N., Mythili, B.: Wild animals intrusion detection using deep learning techniques. Int. J. Pharm. Res. **12**, 1053–1058 (2020)
17. Jocher, G., et al.: YOLOv5 by Ultralytics 2020 (2023). https://github.com/ultralytics/yolov5
18. Matta, V.D., Mudunuri, K.A.V.R.R., Sai Baba, B.C.S., Kiran, K.B., Veenadhari, C.H.L., Prasanthi, B.V.: Single use plastic bottle recognition and classification using Yolo V5 and V8 architectures. In: International Conference on Cognitive Computing and Cyber Physical Systems, pp. 99–106 (2023)
19. Hussain, S.M., prasanthi, b.v., kandula, n., uppalapati, p.j., dasika, s.: enhancing Melanoma Skin Cancer Detection with Machine Learning and Image Processing Techniques. In: Das, P., Begum, S.A., Buyya, R. (eds.) Advanced Computing, Machine Learning, Robotics and Internet Technologies. AMRIT 2023. CCCIS, vol. 1954. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-47221-3_23
20. Cao, C., Chicco, D., Hoffman, M.M.: The MCC-F1 curve: a performance evaluation technique for binary classification. arXiv preprint arXiv:2006.11278. (2020)

21. Christen, P., Hand, D.J., Kirielle, N.: A review of the F-measure: Its history, properties, criticism, and alternatives. ACM Comput. Surv. **56**, 1–24 (2023)
22. Jocher, G., Chaurasia, A., Qiu, J.: YOLO by Ultralytics (Version 8.0. 0) [Computer software]. YOLO by Ultralytics (Version 8.0. 0) [Computer software] (2023)