2024 IEEE Flagship International BIT Conference (BITCON-24)

From Data to Defense: A Deep Learning Approach to Automated Wildlife Monitoring and Anti-Poaching Efforts

Gobinath A
Department of Information Technology
Velammal College of Engineering and Technology
Madurai, India
agn@vcet.ac.in

Jeyakumar N K
Department of Information Technology
Velammal College of Engineering and Technology
Madurai, India
jeyakumar.nkumaran@gmail.com

Gowtham S P
Department of Information Technology
Velammal College of Engineering and Technology
Madurai, India
spgowtham1703@gmail.com

Manjula Devi C
Department of Information Technology
Velammal College of Engineering and Technology
Madurai,India
cmd@vcet.ac.in

Rajeswari P
Department of Electronics and Communication Engineering
Velammal College of Engineering and Technology
Madurai,India
pra@vcet.ac.in

Anandan M
Department of ECE
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
manandan75@gmail.com

2024 International BIT Conference (BITCON) | 979-8-3315-1839-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/BITCON63716.2024.10985534

***Abstract*—Biodiversity has been put in great danger due to the threats of illegal hunting, which endangers many life forms and disrupts ecological balance. The World Wildlife Fund reports that extinction threatens nearly 10000 species every year. The existing methods of monitoring have their merits but will crumble under the size of the wildlife preservation areas and the stealthy operations involved in poaching activities. This project seeks to overcome these problems using state-of-the-art machine-learning techniques such as incorporating YOLOv10 for object detection in real-time and Convolutional Autoencoder (CAE) for the purpose of anomaly detection. We present a new dataset, WLD-72, which consists of 14,400 images from 72 classes, which was developed to train models for wildlife recognition and aggressive human behavior targeted at hiding poaching activities. The system is capable of providing real-time surveillance, high accuracy, and effective anti-poaching deployment which all serve the course of conserving wildlife worldwide. Furthermore, it presents a compromise that can be adopted to different environmental types, thus making it very important in the fight against loss of biological diversity.**

Keywords—YOLO, Computer Vision, Wildlife Conservation, Camera Traps, Object Detection

I. Introduction

Global biodiversity is under a serious threat with wildlife poaching and illegal hunting considered as the greatest wildlife crime by the World Wildlife Fund. According to estimates, 68% of species gone due to anthropogenic reasons since 1970, and these comprise mainly the poaching activities and destruction of habitats. About 10,000 species are estimated to become extinct every year at an alarming pace. To these extremes, we can estimate that poaching particularly that motivated by the unlawful use and commerce of wild animals and plants is expensive ranked is estimated and is between seven billion and twenty-three billion dollars which is also in the top five causes of biodiversity loss related drugs and human trafficking.

Specific species such as elephants, rhinos, tigers, and pangolins are more threatened than others. In the past 10 years, it is estimated that Africa's elephant population has reduced by 30 percent, because of the threat of ivory poaching. Due to poaching, the situation of rhinos and pangolins is similar: over 1 million pangolins have been actively poached out in the last 10 years making them the most trafficked mammal on earth. In India, more than 203 tigers were poached between 2012-2018, despite numerous conservation strategies being put in place itself [1].

Recent developments in technology such as remote sensing, satellite monitoring, and GPS facilitation have improved the monitoring of animals and their movements however, this comes at a cost of data interpretation and the speed of action. Other strategies though important are inadequate in the fight against advanced poachers because of the extensiveness of the geographical region, lack of facilities and infrastructure, and the secretive nature of poaching activities.

Most of the wildlife detection systems require people to examine and identify individuals in the image producing a species classification based on shape, color and size attributes which results into inconsistencies and errors in the classification. However, even though chimpanzees and gorillas, or deer and reindeer appear similar to the naked eye, they behave, live and physically differ. Such accuracy is important to the researcher interested in studying the interrelations and movements of populations over time.

Fig. 1. Challenging animal species (a) Chimpanzee (b) Gorilla (c) Deer (d) Reindeer

As can be observed in Fig. 1, (a) and (b) possess similar body structures but vary in their size and texture of fur while (c) and (d) have similar body forms but differ in shapes of antlers as well as the color of the fur. Such occurrences of overlap in characteristics are major cause of problems in determining the different species. Thus, comprehensive and clear identification information is crucial for improving the effectiveness of species conservation and management, enabling proper surveillance and shielding of poached endangered species from unrepressed illegal hunting.

979-8-3315-1839-4/24/$31.00 ©2024 IEEE

The work's main contributions are the design and formation of a novel Wildlife animal species and Poaching dataset, WLD-72, in the presence of analytical deep learning models which operates ingested data for real-time processing. This novel strategy has high possibilities of improving the accuracy and the speed at which conservation programs are implemented, as it is an effective and flexible for addressing poaching in the world.

Fig. 2. Sample Images of WLD-72 Dataset

II. Related Works

Technology has become one of the most important lines of defence against wildlife poaching due to factors such as crime prediction algorithms and the use of aerial drones in the protection of certain species of animals. Xiaoyuan Yu et al. [2] presented enhancement of sparse coding spatial pyramid matching (ScSPM) for better species classification of wildlife images based on remote camera traps. The method employs descriptor extraction with the help of SIFT and cell-structured LBP and then classification using a support vector machine which has been helpfully working in complicated scenarios of wildlife. E. Bondi et al. [3] discussed the dataset BIRDSAI that was created to track and localize objects of interest in aerial thermal infrared videos long attained by drones in African national parks and game reserves. Also Included are problems like different scales and motion blur. Tsung-Wei Ke et al. [4] during their study focused on the development of a system of aerial photogrammetry concerning wildlife population monitoring and especially taken into account such innovative technique as Mask R-CNN for deep learning based bird detection. In addition to that it has time savings as well on the processing of the results. Nevertheless, the rate of classification could be increased in the following improvements.

In video recordings, Matthias Zeppelzauer [5] has suggested an automated detection algorithm of elephants by deploying a color-based tracking model. This technique also works well in the case of motion occlusions and illumination changes. Meilun Zhou and his colleagues et al. [6] integrate sUAS with deep learning to observe animals within the vicinity of an airport. Their ResNet based model was able to work on species classification with great efficacy, proving the usefulness of sUAS in difficult scenarios. Sathishkumar Samiappan et al. [7] came up with the Aerial Wildlife Image Repository (AWIR), a living archive of animal photography taken by drones which is designed to train AI for monitoring wildlife. H. Nguyen et al. [8] focused on addressing the problem of animal recognition through the application of deep learning, such system is capable of even more enhanced efficiency of monitoring of stoics and managing of the wildlife.

In recent years, researchers, including Christin Carl et al. [9], have attempted to fine-tune a deep learning model which has been trained on classifying European mammals from infrared images. They indicate that additional training on local species could bring about improvement in the detection accuracy. M. Favorskaya et al. [10] have developed a system for detecting animal poses making use of three CNN branches comprising VGG16 and VGG19 and achieving high accuracies although the authors pointed out that the performance could be affected where the training data is not balanced. Suk-Ju Hong et al. [11] evaluated the performance of several deep learning models for bird detection in UAV images and found that the most accurate of all models was Faster R-CNN. Zijing Wu et al. [12] conceived a strategy to locate and count numerous herds of mobile ungulates on satellite images which looks beneficial in the observation of many species across various landscapes.

Challenges in high-level detection methods with the use of UAVs have also been analyzed by Miguel A. Olivares-Mendez et al. [13] where vision sensors were installed on UAVs for herd and poachers detection and in surveillance missions, autonomous control was utilized for remotely piloted control, however, the resolution of the cameras remained a restriction. Drones were also integrated with image processing, machine learning, and robotics by Dane Brown and Daniel Schormann [14] to autonomously find and categorize animals and poachers in the environment where they could tell the difference between poachers and rangers. Stressing on aerial surveillance as a means to counter poaching, Shreya Shivaji Gaikwad et al. [15] adopted YOLO algorithm to identify poachers around the region. To conclude this section, Katie E. Doull et al. [16] assessed the use of drones for poachers spotting and cameras mounted on the drones worked well especially when using thermal cameras at dawn, although there still exists the problem of false exposure detection rate.

III. Proposed Methodology

A. Data Collection and Preparation

The core of the project is to collect and collate the necessary data. This module is specifically aimed at collecting images from various sources i.e. camera traps and drones. These devices are able to collect data from a wide range of inaccessible wildlife regions hence making it possible to cover areas with animal habitation and possible areas of poaching. The data collected is of two main categories:

1) Wildlife Data: This comprises images of various types of animals (for example lions, tigers, elephants etc.) taken in their natural habitats.

2) Poaching Indicators Data: This consists of images that reveal the encroachment of humans in restricted places, pictures of poachers, equipment used for poaching (such as guns, traps, knives), and even motor vehicles used in the act.

B. Data Preprocessing & Annotation

Upon collection of data, it is then cleansed and prepared in a way that makes it suitable for training deep learning models. This module comprises of steps like clean up, resize, and label the images. This procedure enhances the quality of

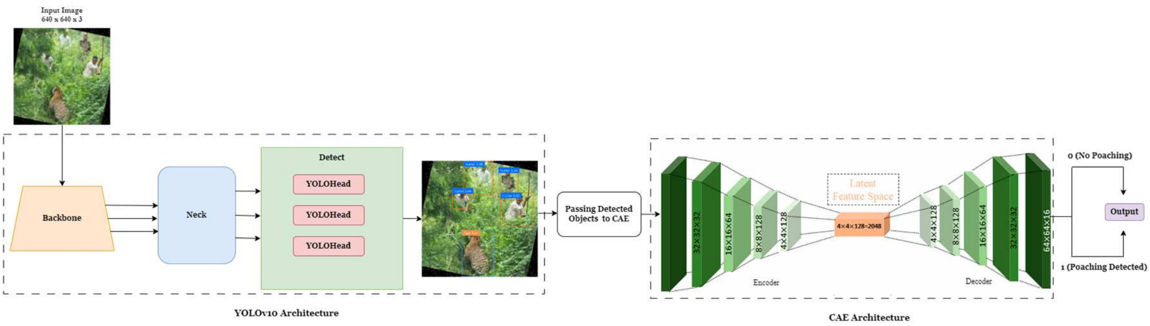

Fig. 3. Architecture Diagram of the Proposed System Integrating YOLOv10 and Convolutional Autoencoder (CAE)

the data eliminating unnecessary noise and paying attention to relevant objects like animals, humans, and weapons. Ensuring standard input to the models, Images are resized to a standard size I_s. For instance:

$$I_s = \frac{I_w}{r} + \frac{I_h}{r} \tag{1}$$

In (1), I_w and I_h denote the image's original dimensions, while r denotes the scaling factor used to standardize the dataset.

The data employed for the current study consisted of 11,923 images in 72 different classes in the beginning which contained many images of animal types from different regions along with images of anti-poaching material. In order to improve the strength of the dataset and make it more diverse, a number of augmentation techniques such as random rotation, switching, scaling, increasing brightness and contrast, and color jittering were employed raising the total count of images to 14,400. Such augmentations serve to reproduce conditions in the real world that contribute to the generalization of the model across different settings. The total dataset had 15,120 labeled examples in 72 categories. Every single image in the dataset was labeled and animals as well as animal poaching indicators such as people, cars, guns etc., were marked using Roboflow. In the course of pre-processing, auto-orientation of images and equal resizing to 640 x 640 pixels by means of a stretch method was done. Such an approach guarantees the production of quality and reasonable data, thus making the dataset ready for purposed model training and evaluation.

C. *Model Development*

The heart of the system will be a real-time object detection model to detect wildlife and poaching activities in real time. As shown in Fig. 3, the structure of the proposed system combines the real-time object detection using YOLO model and Convolutional Autoencoder (CAE) for anomaly detection. The architecture uses the YOLOv10 structure to recognize different objects, including animals and ones that could be poachers, the image is segmented into grid cells and bounding boxes and probabilities for the classes are predicted. YOLO is utilized with the intention of increasing the speed and precision of detection at the same time minimizing the resource consumption.

Following that, the CAE processes the identified objects by reconstructing the input images and measuring the reconstruction errors for the purpose of detecting anomalies, such as the presence of people where there is supposed to be wildlife. This solution is effective in monitoring wildlife and controlling potential threats. The joined framework guarantees simplicity in carrying out all the processes hence improving the efficiency of the system in real time monitoring and detection of risks.

1) Object Detection Using YOLOv10: In order to maximize detection performance on the proposed WLD-72 dataset, YOLOv10 model was treated as the base model. YOLOv10 incorporated a new advanced backbone architecture responsible for feature extraction, thereby increasing the accuracy and performance of the model in detecting objects of interest when monitoring wildlife and detecting poaching. Notably, its design is a dual head model: Three functional heads are employed in the design, with a Classification Head used to categorize every object inside the picture to the correct amenity (this could be a specific type of animal, a person, a car, etc.) and compute the possibility for every class, the second being a Regression Head that encodes information about the spatial coordinates (center, width, and height) of a given object. It defines the loss associated with predicting the bounding box as shown in (2),

$$L_{box} = \lambda_{coord} \sum_{j=0}^{N} \left(\left(x_j - \hat{x}_j\right)^2 + \left(y_j - \hat{y}_j\right)^2 + \left(w_j - \hat{w}_j\right)^2 + \left(h_j - \hat{h}_j\right)^2 \right) \tag{2}$$

In this context, x_j, y_j is the centroid of the bounding box while w_j, h_j are the dimensions of the bounding box length and breadth respectively and also, $\hat{x}_j, \hat{y}_j, \hat{w}_j, \hat{h}_j$ are the respective predicted values.

The loss function which deals with prediction of class is given by (3),

$$L_{class} = -\sum_{j=0}^{N} \left(\left(y_j \log(\hat{y}_j)\right) + (1 - y_j) \log(1 - \hat{y}_j) \right) \tag{3}$$

In (3), the symbol y_j represents the actual label of the class, while the symbol $\hat{y}_j$ represents the predicted class probability.

It is worth mentioning that without non-max suppression, the YOLO version 10 can be used directly even at inference time due to the use of consistent dual assignment strategy

TABLE I. COMPARATIVE ANALYSIS OF YOLO MODELS TRAINED ON WLD-72 DATASET

Model	Parameters	FLOPS (G)	mAP@0.5 (%)	mAP@[0.5:0.95] (%)	Inference Time (ms)	Precision (%)	Recall (%)	IoU (%)
YOLOv7	36.9M	104.7	56.8	47.3	12.5	74.5	71.2	68.5
YOLOv8l	43.7M	165.2	63.4	52.1	14.2	76.1	72.8	70.9
YOLOv9e	58.1M	192.5	70.2	57.5	15.0	79.3	74.6	72.5
YOLOv10-L	24.4M	120.3	**78.2**	**61.4**	16.4	**81.9**	**76.1**	**74**

during training itself. This development makes the post processing straightforward and also minimizes any overheads in the computation enabling real time tracking which is essential in ensuring effective anti-poaching measures. We can write the combined loss of the yolo model in the lines below,

$$L_{total} = L_{box} + L_{class} + L_{objectness} \quad (4)$$

In this context, $L_{objectness}$ refers to the loss incurred by the confidence score with respect to the presence of the object.

2) Object Reconstruction using CAE: The main purpose of the CAE is to acquire effective representations of input data using its encoder-decoder structure. There are two key parts to the CAE: the encoder, which condenses the input data into a lower-dimensional latent space, and the decoder, which then reconstructs the data to its original form. This architecture is particularly well-suited for proposed research work, as it allows for the detection of anomalies in wildlife monitoring by comparing reconstructed images against the original input. In the encoder part, many convolutional layers are implemented, which apply various filters on the input picture and extract important features from it. Each convolutional layer is followed by some nonlinear transformations (ReLU) and max-pooling layers are added, which reduce the size of the obtained feature maps. Such helps in the acquisition of the hierarchical representation of the data, which facilitates the model in learning how to detect the patterns that are useful in recognizing animals and even their would-be hunters. On the other hand, the operation of deconvolutional layers in the decoder is to regenerate the image corresponding to the encoded image only. In this case, deconvolutional layers also known as transposed convolutional layers make it possible to upscale the feature maps back to the input image size. Such symmetric arrangement in both the encoder and decoder permits the CAE to recover the lost content appropriately providing good reconstruction.

a) Latent Space Dimension in CAE: The dimension of the latent space in CAE confers great relevance to the performance of the model. It is the core compressed representation of the input data and choosing an optimal dimension is important to ensure the effectiveness of the model without compromising on the accuracy of the reconstruction of the features. With more capacity, the latent space may express exciting features and characteristics of the data, however, it also risks overgeneralization whereby the representation will learn the variations that are not patterns, but noise. On the other hand, the risk with a configuration with a limited latent space is that even though it may fit well with the available data, it may be incapable of capturing the intricacies of the data.

$$L_{reconstruction} = {}^{1}/_{N} \sum_{i=1}^{N} \left|\left| I_{input}^{(i)} - I_{reconstructed}^{(i)} \right|\right|^{2} \quad (5)$$

Where, N indicates number of samples, $I_{input}^{(i)}$ indicates original input image, $I_{reconstructed}^{(i)}$ indicates reconstructed output image.

A systematic approach was followed to define and refine the latent space that has enabled the CAE to detect anomalies in the wildlife monitoring system effectively. This is aimed at aiding in counter poaching activities. The System Reconstruction Error Threshold will be as shown in (6),

$$Threshold = \mu + k \cdot \sigma \quad (6)$$

Where μ in (6) indicates the average reconstruction error obtained from the training set, σ is the reconstruction error standard deviation, while k represents the number of standard deviations in which the threshold should be set for that particular case (i.e. k=2).

The primary usage of YOLOv10 is identify and localize objects suspiciously involved in poaching like people and vehicles in real time including animals. It gives cropped images of the identified objects, which are then fed to the CAE and classification head for the next step.

D. Model Training

The annotated dataset was divided into: training, validation, and test sets. Then the models were trained on those data.

1) Hyperparameters: Appropriate hyperparameters are employed in the training process in order to improve the performance of the model. Learning rate, batch size, number of epochs, and optimizer are such hyperparameters. Optimally, a learning rate of 0.001 was set in order to achieve a good trade-off between the speed of convergence and stability of the training process. This means that in both cases, the models were able to learn the underlying data distributions without any of them being able to crash the best solution. A batch size of 32 was picked to ensure the available GPU was well utilized without overflowing ounces of memory. The number of epochs was set to 225 which was enough for alterations to be made to the YOLO while for CAE it was limited to 50 epochs so as to extract intricate features and patterns from the dataset. The Adam optimizer was used as it has a learning rate that adjusts itself which is

TABLE II. PERFORMANCE EVALUATION OF CONVOLUTIONAL AUTOENCODER IN ANOMALY DETECTION

S. No	Metric	Value
1	Reconstruction Error (Normal)	0.15
2	Reconstruction Error (Anomalous)	0.35
3	Anomaly Detection Accuracy	91.4%
4	Precision	89.2%
5	Recall	86.9%
6	F1 Score	88.0%
7	Inference Time (ms)	42 ms

very important for effectively training the models especially complicated ones like YOLO and CAE. Dynamic learning rate adjustment as seen in Adam goes a long way in speeding up training and enhancing performance.

2) Loss Functions: YOLO as a model implements a composite loss function which includes objectness loss, classification loss and bounding box regression in order to predict and localize objects accurately. On the other hand, CAE uses Mean Squared Error (MSE) loss to find out how different the reconstruction of a given image is from the original image, meaning that it can also use this to find structural regularities and anomalies in the subject of interest.

The training commences when the annotated dataset is fed into the model. There is an active forward propagation where, the data is passed through the layers of the YOLO and CAE model. At this phase, the systems make predictions and the loss functions measure the predicted outputs accurately as well as the actual outputs. Backward propagation comes in where gradients are found and the model, its weights are changed so that the loss function can lower. This pattern is done over and over again for all the batches and epochs in order to improve the accuracy of the model.

E. Real-Time Monitoring and Alert System

The system that has been designed is intended to facilitate the tracking of the activities of the poachers in real time. In this case the deep learning model will be utilized to process the live feed images allowing the real time detection as described in below.

1) YOLO Object Detection: In the case of YOLOv10, first the image is fed through the network and locations of objects such as animals, people, cars and poaching related objects are detected in the form of bounding boxes. For each of these objects that have been detected, the coordinates of the bounding box, its class, and a confidence score of its presence are presented from the model.

2) Passing Detected Objects to CAE: Individual detected objects from the source image are extracted using the bounding boxes developed by YOLO. These extracted occluded object images are then forwarded to the CAE for analysis.

3) CAE Anomaly Detection: CAE attempt to reconstruct each of the cropped-out object images and perform a comparison with the original sub-image. Reconstruction error will be computed for each of the objects that the CAE assesses to be normal (for example, wildlife) or anomalous (for example, human, vehicles). If a human or vehicle is detected by YOLO, and the CAE gives a high reconstruction error, then it implies there is an unusual behavior going on. Because the CAE relies on reconstruction error, it can also present a output flag; 0 (Normal Activity) – No anomalies found within detection, 1 (Abnormal Activity) – Presence of suspicious people or things, which requires reaction or warning.

4) Final Decision and Alert Generation: The system will issue an alert when the CAE marks any sighted object as abnormal, and this alert will comprise details such as the class of object, location information, and the level of abnormality exhibited by the object. The alert shall also contain precise geolocation coordinates and time of the occurrence of the poaching event. This enables rangers to act quickly, increasing the likelihood of averting illegal activities.

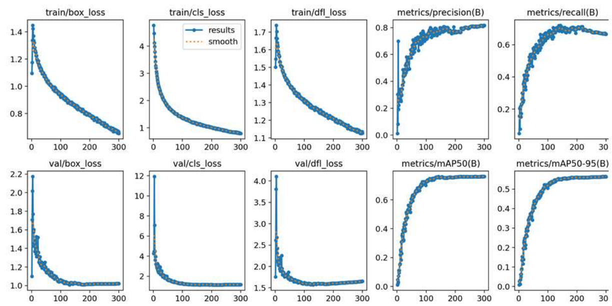

Fig. 4. Training Graphs for Proposed Model Performance Metrics

IV. RESULTS AND DISCUSSION

The model that has been suggested uses the capabilities of YOLOv10 for object detection in conjunction with a CAE for the purpose of anomaly detection and such a system works effectively in the conservation of wildlife and in detection of poaching instances as well. Essential performance measures including precision, recall, accuracy and loss values as a function of the training epochs are shown in Fig. 4. This helps in appreciating more the tendencies of the model during learning and how effective it is in detecting wild animals and poaching respectively.

A. Performance Evaluation of Proposed System

1) YOLOv10: As seen in Table I, the performance evaluation of YOLOv10 in terms of mAP achieved 78.2% at 0.5 IoU while 61.4% was achieved for the more specialized mAP@[0.5:0.95] metrics, placing it on the top among the other segmentation models. The high precision (81.9%) and recall rate (76.1%) render its ability to work even in real time for detecting wildlife as well as indicators of possible poaching such as vehicles and weapons the best. YOLOv10 has the longest inference time (16.8 ms) compared to all the others at most, but it provides the best compromise between accuracy and speed which means the system can monitor and generate an alert in real-time satisfactorily. It is therefore the most appropriate model for use in this proposed study since effective response to poaching and other anomalies for that matter calls for quick detection of such threats.

2) CAE: As presented in Table II, the CAE demonstrated a low reconstruction error of 0.15 when reconstructing images of only the detected wildlife, indicating a good ability in recognizing the normal behavior. In comparison, objects that were classified as anomalies (like people and cars) produced a higher reconstruction error of 0.35, thus validating the CAE's efficacy in recognizing abnormal behavior.

Fig. 5. The system identified suspicious activity involving unauthorized human presence and vehicles flagged as poaching

B. *Alert Generation*

In order to communicate effectively with the respective forest authorities concerning the poachers or system-induced irregularities, the incident relating to which happened, a Python Twilio library was implemented. This enabled the creation of automated recorded voice calls and SMS notifications to the relevant personnel, accordingly ensuring that such personnel are alerted in time. Moreover, Smtplib and Email Libraries were used to create email alert notifications for the system which makes it an alert system adding on the efficiency of the response of the authorities.

V. Conclusion

This research work demonstrates great potential for addressing issues related to wildlife conservation by incorporating high-end machine learning models which can geolocate a specific animal or illegal activity. The system showed stable outcomes by developing the WLD-72 dataset of animal and poaching relevant objects. YOLOv10 reached a mAP of 78.2%, with precision of 81.9% and recall of 76.1% showing its ability in object detection. Another layer of anomaly detection was provided with the CAE, which was able to identify and raise the alarm on wildlife crimes. If such a system is established where automatic alerts would be sent to the relevant authorities in real time, it could greatly enhance the fight against poaching and the preservation of rare species. In conclusion, this study offers a cost-effective solution to one of the main problems, the vastness of space in monitoring and protecting wildlife, which is a great milestone in the use of technology for conservation purposes.

In future developments, the system's resilience will be improved in order to withstand extreme peripheral conditions such as operating optimally in low light conditions, at night and in poor visibility weather conditions such as rain, fog, and thick mist. To achieve proper wildlife detection and poaching activity monitoring even in these harsh conditions, advanced technologies including thermal imaging, infrared cameras, and specific improved imaging techniques will be investigated. This will expand the application of the system's capabilities, making it indispensable in all environmental monitoring without fail.

References

[1] https://projecttiger.ntca.gov.in/tiger-mortality/

[2] Yu, X., Wang, J., Kays, R. et al., Automated identification of animal species in camera trap images. J Image Video Proc 2013, 52 (2013). https://doi.org/10.1186/1687-5281-2013-52

[3] E. Bondi et al.,, "BIRDSAI: A Dataset for Detection and Tracking in Aerial Thermal Infrared Videos," 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass, CO, USA, 2020, pp. 1736-1745, doi: 10.1109/WACV45572.2020.9093284.

[4] Ke T-W, Yu SX, Koneff MD, Fronczak DL, Fara LJ, Harrison TJ, et al., (2024) Deep learning workflow to support in-flight processing of digital aerial imagery for wildlife population surveys. PLoS ONE 19(4): e0288121. https://doi.org/10.1371/journal.pone.0288121

[5] Zeppelzauer, M. Automated detection of elephants in wildlife video. J Image Video Proc 2013, 46 (2013). https://doi.org/10.1186/1687-5281-2013-46

[6] Zhou, Meilun, et al., "Improving animal monitoring using small unmanned aircraft systems (sUAS) and deep learning networks." Sensors 21.17 (2021): 5697.

[7] Kissling, W. Daniel, et al., "Development of a cost-efficient automated wildlife camera network in a European Natura 2000 site." Basic and Applied Ecology (2024).

[8] Nguyen, H., Maclagan, S. J., Nguyen, T. D., Nguyen, T., Flemons, P., Andrews, K., Phung, D. "Animal Recognition and Identification with Deep Convolutional Neural Networks for Automated Wildlife Monitoring," 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Tokyo, Japan, 2017, pp. 40-49, doi: 10.1109/DSAA.2017.31.

[9] C. Carl, F. Schönfeld, I. Profft, A. Klamm, and D. Landgraf, "Automated detection of European wild mammal species in camera trap images with an existing and pre-trained computer vision model," Eur. J. Wildl. Res., vol. 66, no. 4, pp. 1–7, Aug. 2020, doi: 10.1007/S10344-020-01404-Y

[10] M. Favorskaya and A. Pakhirka, "Animal species recognition in the wildlife based on muzzle and shape features using joint CNN," Procedia Comput. Sci., vol. 159, pp. 933–942, Jan. 2019, doi: 10.1016/J.PROCS.2019.09.260.

[11] Hong, Suk-Ju, et al. "Application of deep-learning methods to bird detection using unmanned aerial vehicle imagery." Sensors 19.7 (2019): 1651.

[12] Wu, Zijing, et al. "Deep learning enables satellite-based monitoring of large populations of terrestrial mammals across heterogeneous landscape." Nature communications 14.1 (2023): 3072.

[13] Olivares-Mendez, Miguel A., et al. "Towards an autonomous vision-based unmanned aerial system against wildlife poachers." Sensors 15.12 (2015): 31362-31391.

[14] D. L. Brown, D. Brown, and D. Schormann, "Poacher detection and wildlife counting system," Accessed: Apr. 02, 2022. [Online]. Available: https://www.researchgate.net/publication/335378767

[15] Shreya Shivaji Gaikwad, Pranit Mohan Mhatre, Cyril Silvester Dabre, Ms. Nazneen Ansari, "Poacher Detection using YOLO Algorithm." International Journal of Engineering Research & Technology (2021).

[16] Katie E. Doull, Carl Chalmers, Paul Fergus, Steve Longmore , Alex K. Piel and Serge A. Wich "An evaluation of the factors affecting 'poacher'detection with drones and the efficacy of machine-learning for detection." Sensors 21.12 (2021): 4074. https://doi.org/10.3390/s21124074