Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)
VOLUME:8, ISSUE:1, FEBRUARY 2025, PAPER NUMBER:47

Monitoring the Moment of Wild Animals an Generating an Alert System Using Deep Learning Algorithm

*Note: Sub-titles are not captured in Xplore and should not be used

1st Gopari Vivek
Dept. of Computer Science and Engineering
Chaitanya Bharathi Institute of Technology
Hyderabad, India
vivekgopari@gmail.com

2nd Vadla Rahul
Dept. of Computer Science and Engineering
Chaitanya Bharathi Institute of Technology
Hyderabad, India
rahulchary902@gmail.com

3rd Dr M.Subramaniam
Dept. of Computer Science and Engineering
Chaitanya Bharathi Institute of Technology
Hyderabad, India
msubramaniam_cse@cbit.ac.in

Abstract—To support conservation and management choices, effective and trustworthy wild animal monitoring in their natural environments is crucial. Automatic hidden cameras, often known as "camera traps," are becoming a more and more common technique for wildlife monitoring because of its efficacy and dependability in gathering data on wildlife in a big volume, continually, and without causing any inconvenience. Nevertheless, manually processing such a vast number of photos and movies taken with camera traps is very costly, time-consuming, and tedious. For ecologists and scientists, this poses a significant challenge to their ability to track wildlife in its natural habitat. In this study, we present a framework to develop automated animal detection in the wild, with the goal of developing an automated wildlife monitoring system, by utilising recent advancements in deep learning techniques in computer vision. More specifically, we train a computational system that can automatically identify species in animal photos by filtering them using a single-labeled dataset from the citizen scientist-run Wildlife Spotter project and the most advanced deep convolutional neural network architectures.

Index Terms—component, formatting, style, styling, insert

I. Introduction

One of the main tasks in ecology is to observe wild animals in their natural settings. Earth's ecosystems are undergoing quick, unique, and significant changes due to overexploitation of natural resources brought on by the world's population expansion and unrelenting pursuit of economic development. Human activity has changed a growing portion of the land surface, which has affected the number, habitat, and behaviour of wildlife. Seriously, a great number of wild species on Earth have been exterminated, and a great number of species are brought into new environments where they have the potential to cause havoc on human and natural systems. For this reason, tracking wild animals is crucial because it gives scientists data to support conservation and management choices that

Identify applicable funding agency here. If none, delete this.

keep ecosystems diversified, stable, and resilient in the face of change. In order to monitor wild animals, a number of contemporary technologies have been developed. These technologies include motion-sensitive camera traps, satellite and global positioning system (GPS) tracking, wireless sensor networks, radio tracking, and more. Motion-triggered remote cameras, sometimes known as ”camera traps,” are becoming a more and more common tool for wildlife monitoring because of their unique characteristics, increased commercial availability, and simplicity of setup and use. For example, a standard covert camera model may record high-definition photographs during the day and at night. It can also gather temperature, time, and moon phase data that is integrated with the im-age data. Furthermore, wide-ranging and adjustable camera settings enable continuous and covert tracking of animals. A camera may record thousands of pictures in a row after it is completely charged, yielding a significant amount of data. since of these features, camera traps are an extremely useful tool for ecologists since they allow them to record every facet of animals. If captured properly, visual data is a wealth of information that can be used by scientists to gather evidence for answering questions related to ecology. Some of these questions include: what are the spatial distributions of rare animals? Which species, like bandicoots, are threatened and require protection? Which group of pest species, like red fox and rabbit, need to be controlled? These are just a few examples of important questions to comprehend the populations of wild animals, their ecological relationships, and population dynamics. To this aim, setting up multiple camera traps in the field to gather picture data of wild animals in their natural surroundings has become a newly popular method employed by ecologists. Thanks to advancements in digital technology, camera trapping is becoming more and more popular for monitoring wildlife. Modern camera traps come with automated system components and lower purchase costs, but the analysis of large amounts of camera trap images has traditionally been done by hand. Even though the human visual system can interpret images quickly and easily, it is quite costly to process such a large number of images by hand. For instance, from 2010 to 2013, 225 camera traps located around Tanzania's Serengeti National Park allowed the Snapshot Serengeti project1 to collect 3.2 million photos. Mil-lions of images of wildlife taken in Australia's dry rangelands and tropical rainforests were gathered by Wildlife Spotter2, a related effort. Regretfully, most acquired photographs are difficult to interpret, even for humans, because of the automatic method used to snap the trap cameras. Only a small portion of the gathered photos are in good enough shape.

Many photos only show a portion of the animal item's body; in other photos, the animal object is photographed in its whole but is too far away from the camera (Figure 2b), or it is partially obscured. Additionally, a significant portion of the images are in grayscale due to the fact that they were taken at night with the assistance of an infrared flash.

Our goal in this research is to develop a fully autonomous wildlife spotting system by creating a framework for animal recognition in the wild. The state-of-the-art capabilities of current deep CNN models for image classification, particularly the proof that automated recognition can outperform humans on some object recognition tasks in the Image Net competition, serve as our driving force. We conduct experiments using datasets from the Wildlife Spotter project, which includes a substantial number of photos captured by trap cameras installed by Australian scientists. More precisely, we split the wild animal identification automation into two more jobs because the Wildlife Spotter dataset contains both animal and non-animal images: (1) Based on the assumption that there would be animals in the photographs, wildlife detec-tion—which is essentially a binary classifier—can categorise input images into two groups: ”animal” or ”no animal”. Additionally, a multiclass classifier for wildlife identification is used to assign a specific species to each input image con-taining images of animals. Fundamentally, every assignment involves a deep CNN-based classifier that is trained using pre-prepared datasets that volunteers have manually labelled. To create the framework for comparisons, a number of carefully chosen deep CNN designs are used. If Task 1 is successful, a significant amount of non-animal image data that citizen annotators are now wasting their time on will be automatically filtered away, enhancing the effectiveness of citizen science-based programmes (like Wildlife Spotter). This method can save a significant amount of time and money, as demonstrated by our testing results on the Wildlife Spotter datasets. Because of this, the main contribution of this work is that deep learning could be used to develop a fully automatic image classification system on a large scale with enough data and computing infrastructure, relieving scientists of the laborious task of manually processing millions of images—a task that project managers believe computers simply cannot perform. Furthermore, by merging our suggested framework with the current citizen science initiative, a ”hybrid” image classifier with an automatic recommendation system that offers volun-teers insightful ideas to expedite their classification decisions is created.

II. LITERATURE REVIEW

1) A.Sathesh et al 2022 [1] The YOLO method is used by the suggested system to find predefined items in the field. The hardware's camera captures the object and transfers the picture to a server when a match is discovered. The farmer receives the collected image via email from the Intelligent Surveillance System. A buzzer that the farmer can regulate is also automatically activated by the system. This system relies heavily on the AIOPENCV framework, which also sounds the buzzer. Before the acquired image is used to train the model, it is compressed and subjected to preprocessing. To ensure dependable and real-time performance, the training process entails feature extraction, extracting the required patterns from the image, and then feature fusion and dimension reduction.

2) Shubham Mishra et al 2022[2] Pandas, OpenCV, and Python were used in the system's implementation. It entails gaining access to the camera and installing the required libraries. By applying thresholding and comparing frames, the system detects motion. Rectangles are drawn on the frames using the coordinates of the contours, which are used to identify objects. Timestamps of object entry and exit events are recorded by the system. The gathered information is kept in a pandas data frame and may be exported to a CSV file for additional examination. Real-time monitoring and object detection in front of the camera are provided by the system.

3) D. Ranparia et al 2021 [3] A camera, a number of sensors, and an Arduino board serve as the main components of the suggested system. The camera starts to take pictures and record a video for a few minutes when it senses motion within a 10-meter radius. Onboard and cloud storage are used to store the picture and video. Furthermore, a SIM900A module is used to automatically send a message to a registered number with information about intrusions as well as temperature and hu-midity readings from a DHT11 sensor. While unauthorised in-dividuals without RFID tags are processed further, authorised staff with valid RFID tags have their attendance automatically recorded. The system detects items and distinguishes between human and animal invaders using Cascade Classifiers based on Haar features. An alarm is set out to alert people to the intrusion if a human is found. The number of PIR sensors that are triggered by an animal intruder is used by the system to identify the proper course of action. These sensors are mounted vertically on nearby watchtowers, fences, poles, and access points.

4) Mohit Korche et al 2021 [4] The suggested system uses PIR sensors to identify the position of the animals and LDRs positioned vertically to identify their size when detecting their presence in the agricultural area. The APR board detects the animal and sounds an alert to reroute it. A message is sent to the farmer and a flash light is turned on throughout the night. The animal's presence and the LDR readings are shown on an

LCD panel. The farmer receives a warning message about the intrusion via a GSM module.

5) Yadahalli et al. 2020 [5] The suggested system makes use of an Arduino Uno board that is linked to a number of cameras and sensors. The device triggers the camera, takes a picture of the intruder, and shows it on a TFT display when the PIR sensors detect motion within a 10-meter radius. A GSM module is used to automatically generate and send a message with sensor readings and details about the intrusion to the owner's registered phone number. The farmer has the ability to activate or disable the system for access control. When the PIR sensors pick up human movement, the owner notification system and camera are activated. The farmer receives visual information from the TFT display and text messages from the GSM module. For audio alerts, there is an alarm or buzzer incorporated. The size of the trespassing animal is determined by the number of triggered PIR sensors, which aids in determining the best course of action to safeguard the crops. The technology notifies the owner and those in the vicinity of the intrusion using audio alerts, text message notifications, and image transmission.

III. EXISTING SYSTEM

Currently, the Fourier transform is used for object detection and segmentation in animal recognition systems. Elephants have been known to leave their normal habitats in search of food and resources due to deforestation, which has led to conflicts with people and damage to cultivated areas. Numerous techniques, such as junction boxes, alarm systems, and crop protection based on acoustic signals, have been proposed to keep animals away from crops. These systems are now in use, although they can be costly, complicated, and contribute to noise pollution. Despite these initiatives, there is still a need for a practical and efficient crop protection solution. It is necessary to create an intelligent system that can protect farms and animals alike.

IV. METHODOLOGY

4.1. Convolutional Neural Network

Convolutional neural networks, or CNNs or ConvNets, are a subclass of neural networks that are particularly good at processing input with a topology resembling a grid, like images. A binary representation of visual data is what makes up a digital image. It consists of a grid-like arrangement of pixels with pixel values to indicate the colour and brightness of each pixel when it is implemented. The moment we perceive an image, the human brain analyses a tremendous quantity of data. Every neuron functions within its own receptive area and is interconnected with other neurons to include the full visual field. Similar to how individual neurons in a biological vision system respond to stimuli exclusively within the limited area of the visual field known as the receptive field, each neuron in a CNN analyses information exclusively within its receptive field. Layers are stacked so that a CNN is used to identify more complicated patterns (faces, objects, etc.) after simpler patterns (lines, curves, etc.) are detected. A type of deep neural network used in deep learning that is most frequently used to analyse visual imagery is the convolutional neural network. Due to the shared-weight architecture of the convolution kernels, which shift over input features and produce translation equivariant responses, they are sometimes referred to as shift invariant or space invariant artificial neural networks (SIANN). The majority of convolutional neural networks are, however, merely equivariant to translation rather than invariant. The fields in which they find use include financial time series, natural language processing, recommender systems, medical image analysis, image and video recognition, image classification, image segmentation, and natural language processing. Because of how much the connectivity pattern between neurons in convolutional networks mimics the structure of the animal visual cortex, these networks were inspired by biological processes. The receptive field, a small portion of the visual field where individual cortical neurons are solely responsive to inputs, is known as this. The complete visual field is covered by the partial overlap of the receptive fields of distinct neurons. Structure of Convolutional Neural Nets A convolutional layer, a pooling layer, and a fully connected layer are the three layers that make up a CNN.

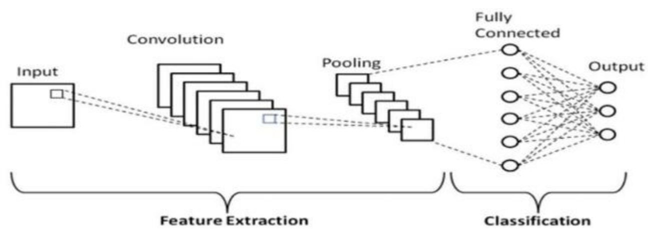

Fig. 1. Fig1.Architecture of CNN

4.2. Revisiting YOLOv5 algorithm

The YOLO algorithm's feature is its ability to maintain a specific accuracy rate while operating at a rapid speed, and it has a low likelihood of misidentifying background objects as recognised objects[6]. It has great adaptability and can provide accurate classification results for a wide range of objects. The YOLO series algorithm's fundamental concept is to split an image into many identically sized grids [6]. Each grid has multiple characteristics, such as the location (x and y) of the object to be detected, the form (length and breadth) of the candidate box, the probability level, and the type of object to be identified. Those grids will be in charge of predicting one target that falls within the grid.

4.3. Proposed System

We introduce our suggested system for classifying images, along with its implementation on the Wildlife video datasets. We begin by describing the datasets. Next, we present a paradigm for identifying animals that is based on CNN. We must first gather the data video, or input video, of the wild animal. After that, we build two environments in which to test our suggested framework for the duties of identifying and detecting wildlife. We concentrate on using the most recent state-of-the-art CNN architectures for both detection and recognition in this work because it has been demonstrated that CNNs perform better than alternative methods in the field of picture categorization. Lastly, we describe a few of the CNN designs that we used in our implementations and experiments.

ADVANTAGES Decisions for management and conserva-tion of wild animals must be based on accurate and efficient monitoring of these creatures in their natural environments. Important feature identification can be done without human supervision. Its image identification and categorization accu-racy is very high.

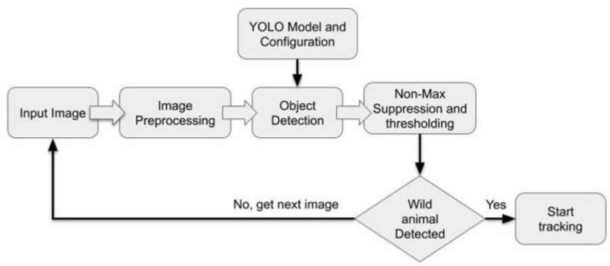

Fig. 2. The flow chart of capturing image and classifying it using yolo technique

V. EXPERIMENTAL RESULTS

A camera detects animals in the wild, records the animal, and uploads the image to a server. The animal is then cat-egorised by the model using its photo, allowing for better monitoring. An alert message is sent to the forest officials' and farmer's email addresses, along with the taken photograph. The technology has a 94% accuracy rate in identifying ani-mals.

A. Detecting Image of Tiger and Elephant

The Fig. 3 and Fig. 4 represents the animal's Tiger and Elephant

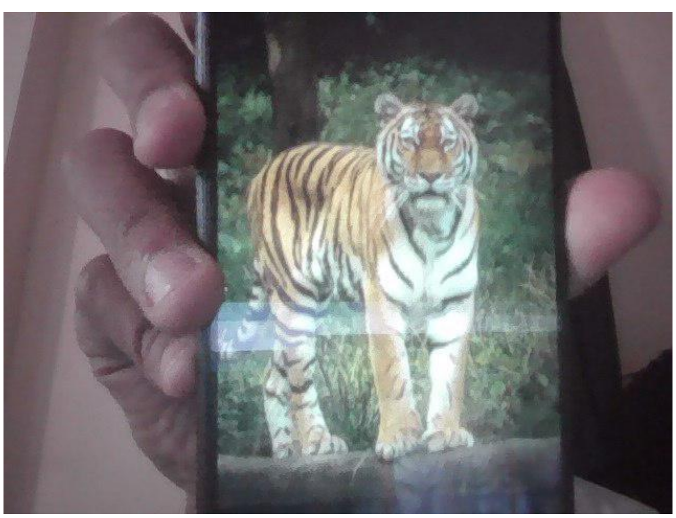

Fig. 3. Detected Tiger Image

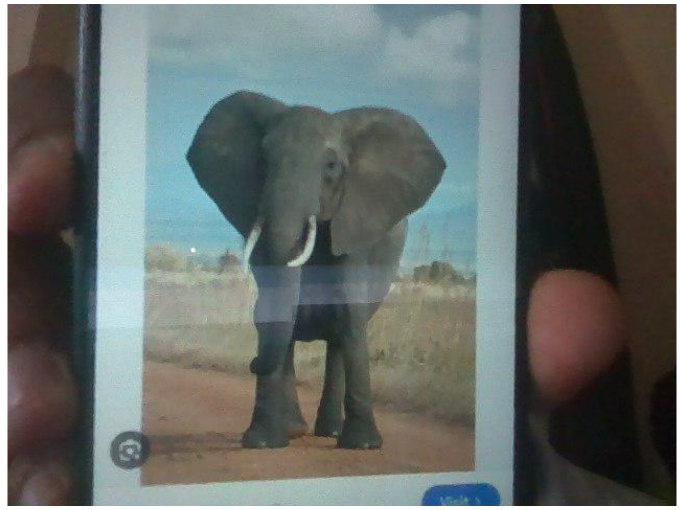

Fig. 4. Detected Elephant Image

The mean squared error (MSE)loss is the loss function employed in the model. During training, a drop in both the IOU loss and total loss indicates that the model is becoming more accurate in identifying items. Reduced losses signify increased object detection accuracy.

The following Fig. 5 and Fig. 6 shows the IOU loss and total loss arrived for Animal classification.

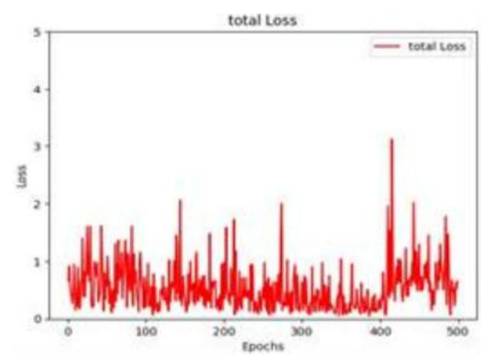

Fig. 5. Total loss

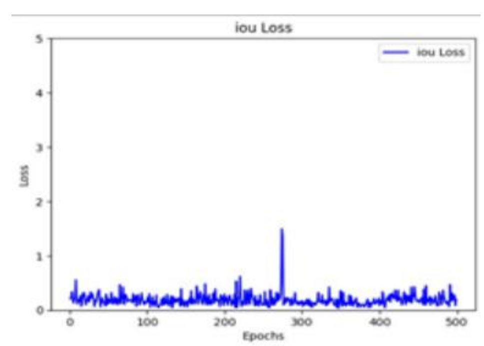

Fig. 6. IOU loss

CONCLUSION

In this research, we presented and illustrated the viability of a deep learning method towards building a scalable automated wildlife monitoring system using the Wildlife Spotter dataset, which contains a huge number of photos captured by trap cameras in South-central Victoria, Australia. Our models identified the three most common animals (bird, rat, and bandicoot) with over 96% accuracy and around 90% accuracy in photos including animals. Furthermore, the system has proven to be reliable, stable, and appropriate for handling photos taken in the field when tested in various balanced and imbalanced experimental setups. We are investigating several approaches to boost the system's efficiency, such as expanding the dataset, using more complex CNN models, and taking advantage of particular characteristics of camera trap photos. In order to address the issue of extremely imbalanced data, we would look into transfer learning as a means of developing a completely automated wild animal recognition system. Our immediate goal is to create a "hybrid" wild animal classification framework, with an automated module that serves as a recommendation system for the Wildlife Spotter project, which is now dependent on citizen research.

REFERENCES

[1] Natarajan B, R Elakkiya, R Bhuvaneswari, Kashif Saleem, Dharminder Chaudhary. "Creating Alert Messages Based on Wild Animal Activity Detection Using Hybrid Deep Neural Networks and Telegram Integration." In Proceedings of the International Conference on Artificial Intelligence and Wildlife Conservation,2024. [2] A. Nidhi, B. Sai Krishna, B. Aditya, Sk Afeef Ur Rehman, Dr. Ravi Mathey. "Wild Animal Detection and Alert System Using YOLOv8: Integrating Telegram Notifications." In Proceedings of the International Conference on Wildlife Monitoring and Conservation,2023.
[3] C.S. Madhumathi, V. Naveen, M. Sanjay Kumar, M. Mohammed Aslam. "Advanced Wild Animal Detection and Alert System using YOLO V5 Model",2023. [4] K. Rajitha and P. Seema, Animal Detection and Warning System using Faster regions with convolutional neural networks (R-CNN), 2022. [5] Ding Ma, Jun Yang. "YOLO-Animal: An efficient wildlife detection network based on improved YOLOv5", 2023. [6] A. Gomez, A. Salazar, F. Vargas. "Towards Au-tomatic Wild Animal Monitoring: Identification of Animal Species in Camera-Trap Images Using Very Deep Convo-lutional Neural Networks and Real-time Telegram Alerts." In Proceedings of the International Conference on Computer Vision and Wildlife Conservation, 2022. [7] M Tan, W Chao, J K Cheng et al., "Animal Detection and Classification from Camera Trap Images Using Different Mainstream Object De-tection Architectures[J]", Animals, vol. 12, no. 15, 2022. [8] A Verma, V Sangwan and N. Shukla, "Wild animal species detection using deep convolution neural network[M]", Recent Trends in Communication and Electronics, pp. 406-410, 2021.