

# AUTOMATED IDENTIFICATION OF WILDLIFE SPECIES VIA DEEP LEARNING MODULE

R.Senthil Kumar<sup>1</sup>, S.Saranya<sup>2</sup>, R.Sivasankari<sup>2</sup>, S.Yazhini<sup>2</sup>

Assistant Professor<sup>1</sup>, Final Year Student<sup>2</sup>

Department of Electronics and Communication Engineering  
SSM Institute of Engineering and Technology Dindigul, Tamilnadu, India

## ABSTRACT:

*In recent years, rapid advancements in artificial intelligence and computer vision have significantly transformed monitoring systems across various application domains. Wildlife monitoring is an essential task in forest conservation, biodiversity protection, and prevention of human-wildlife conflict. Traditional wildlife observation methods rely on manual monitoring and analysis of camera trap images, which is time-consuming and inefficient under both daytime and night-time conditions. To overcome these challenges, this project proposes an automated wildlife detection system based on deep learning techniques. The system utilizes the YOLOv11 object detection model to automatically detect and classify animals from both daytime visible-light images and night-time infrared images captured using camera traps. Image preprocessing and dataset augmentation techniques are applied to improve detection accuracy under varying illumination and environmental conditions. The proposed system identifies wildlife species, localizes them using bounding boxes, and reduces the need for continuous human supervision. By automating the wildlife detection process, the system enhances monitoring efficiency, supports forest surveillance, and aids in wildlife conservation efforts during both day and night.*

**Keywords:** *Wildlife Detection, YOLOv11, Deep Learning, Infrared and Daylight Images, Object Detection, Automated Monitoring.*

## I. INTRODUCTION

Wildlife plays a vital role in maintaining ecological balance and biodiversity in natural ecosystems. Rapid urbanization, deforestation, and climate change have significantly increased interactions between humans and wildlife, leading to habitat loss and rising human-wildlife conflicts. Continuous monitoring of wildlife is essential for effective conservation, forest management, and protection of endangered species. Traditional wildlife monitoring methods involve manual observation and physical surveys, which are labor-intensive, time-consuming, and often inaccurate. With the advancement of camera trap technology, large volumes of wildlife images are captured daily during both daytime and night-time conditions. However, manual analysis of these images becomes difficult and inefficient.

In day-to-day forest monitoring activities, forest officials and researchers are often engaged in multiple tasks, leading to delayed analysis and reduced awareness of wildlife movement patterns. Many wild animals exhibit nocturnal behavior, while others are active during daylight hours, making continuous monitoring necessary. Monitoring wildlife using both visible-light images during daytime and infrared images during night-time presents challenges such as illumination variations, noise, poor contrast, and background complexity. Without automated intervention, identifying animals from such images becomes difficult and prone to errors. Recent advances in deep learning and computer vision have enabled automated wildlife detection with improved accuracy and reliability.

To address these challenges, this project focuses on developing an automated wildlife detection system using the YOLOv11 deep learning model. The system detects animals from both daytime and night-time images and accurately localizes them using bounding boxes. By automating the detection process, the proposed system reduces manual effort, improves monitoring efficiency, and supports effective wildlife conservation and forest surveillance.

## II. LITERATURE SURVEY

In the year 2016, a research work titled “*You Only Look Once: Unified, Real-Time Object Detection*” was proposed by Joseph Redmon et al., which introduced the YOLO object detection framework. This work focused on real-time object detection by treating detection as a single regression problem. The model demonstrated high speed and good accuracy, making it suitable for real-time applications such as surveillance, traffic monitoring, and environmental observation systems.

Later in 2018, Norouzzadeh et al. proposed a deep learning-based approach for automatically identifying and classifying wild animals in camera trap images. The system utilized convolutional neural networks to detect animal presence and species from large-scale datasets collected during both day-time and night-time. Although the approach significantly reduced manual effort in wildlife monitoring, it faced challenges under low-light and infrared imaging conditions.

In the year 2019, Beery et al. presented a wildlife recognition system using camera trap datasets captured under diverse environmental conditions. The study focused on addressing issues such as occlusion, background clutter, and illumination variations across day and night scenarios. While the detection accuracy was improved, the system required further optimization for real-time deployment.

In the proposed system, an automated wildlife detection model based on the YOLOv11 framework is developed. The system focuses on detecting animals from both daytime visible-light images and night-time infrared images with improved accuracy and faster processing speed. By overcoming limitations of existing approaches, the proposed system enhances wildlife monitoring efficiency and supports effective conservation efforts.

### III.BLOCK DIAGRAM

The main platform of this project is a deep learning-based object detection system developed using the YOLOv11 framework. The system operates on computer vision principles and is designed to automatically detect and classify wildlife from images captured in natural environments. YOLOv11 is an efficient and high-speed object detection model that enables accurate real-time detection under varying lighting conditions.

The primary input to the system consists of images collected from camera trap devices deployed in forest regions. These images include both daytime visible-light images and night-time infrared images. The captured images are preprocessed using techniques such as resizing, normalization, and noise reduction to enhance image quality. Dataset augmentation is also performed to improve model robustness and generalization.

The preprocessed images are then fed into the YOLOv11 detection model, which extracts features and performs object detection. The model identifies animals present in the images and localizes them using bounding boxes. Classification is carried out to distinguish different wildlife species based on trained labels, without requiring continuous human intervention.

The final output of the system is an annotated image displaying detected animals along with their class labels and bounding boxes. This automated workflow improves monitoring efficiency, reduces manual effort, and supports effective wildlife surveillance and conservation activities.

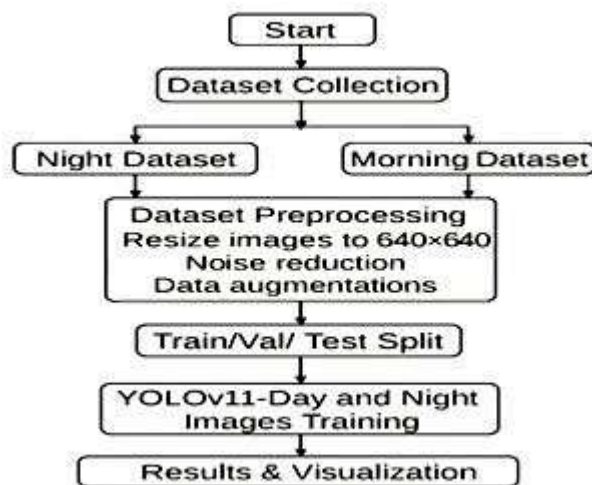


Fig 1: Block Diagram of Automated Wildlife Detection System

## 1. Camera Trap Dataset

Camera trap images serve as the primary data source for the proposed wildlife detection system. These camera traps are strategically deployed in forest and wildlife reserve areas to observe animal activity without causing disturbance to the natural ecosystem. The cameras operate in a non-invasive manner, allowing wildlife to move freely in their habitat. They are motion-triggered devices that automatically capture images whenever movement is detected within their sensing range. This mechanism ensures continuous and autonomous data collection without the need for human supervision. As a result, large volumes of wildlife images can be gathered efficiently over extended periods.

The dataset consists of both daytime and nighttime images captured under real-world environmental conditions. Daytime images are acquired using visible-light sensors, providing clear visual details such as color, texture, and background context. Nighttime images are captured using infrared imaging technology, enabling the detection of nocturnal animals under low-light or no-light conditions. The inclusion of infrared images ensures effective monitoring of wildlife activity during nighttime hours. By incorporating images from different times of the day, the dataset supports round-the-clock wildlife surveillance.

Additionally, the dataset contains images captured under varying environmental conditions such as changes in weather, vegetation density, and illumination. This diversity enhances the robustness of the detection system. The availability of diverse image samples reduces bias and improves the generalization capability of the detection model. Overall, the camera trap dataset plays a crucial role in enabling accurate, reliable, and continuous wildlife monitoring across different forest environments.

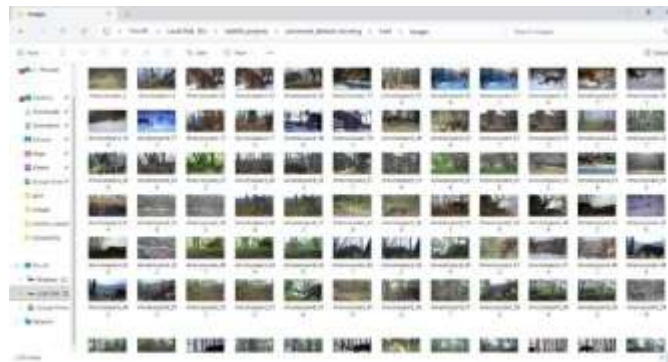


Fig 2: Day-Time Dataset Images

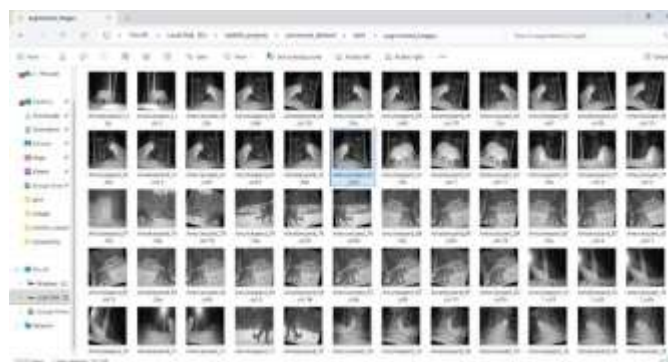


Fig 3: Night-Time Dataset Images

## 2. Image Preprocessing and Annotation

Image preprocessing is an essential stage in the proposed wildlife detection system to improve detection accuracy and model reliability. Raw camera trap images often suffer from challenges such as noise, illumination variation, motion blur, and complex background clutter. These factors can negatively impact the performance of object detection models. To address these issues, preprocessing techniques are applied to enhance image quality and standardize the dataset. Image resizing is performed to ensure that all input images conform to the fixed input dimensions required by the YOLOv11 model. This step helps maintain consistency during training and inference.

Normalization is applied to standardize pixel intensity values across the dataset. This process reduces variations caused by lighting differences between day and night images. Noise reduction techniques are also employed to improve image clarity and suppress

irrelevant background information. These preprocessing steps help the model focus on meaningful visual features related to animals rather than environmental noise. Through preprocessing, the overall quality and usability of the dataset are significantly improved.

Following preprocessing, image annotation is carried out to prepare the dataset for supervised learning. Annotation involves drawing bounding boxes around animals present in the images and assigning appropriate class labels. These labeled bounding boxes provide precise spatial and semantic information required for training the detection model. Accurate annotation enables the model to learn object boundaries, spatial relationships, and class-level characteristics effectively. As a result, the annotated dataset supports improved localization accuracy and reliable classification performance during wildlife detection.



**Fig 4:** Annotated Dataset Images

### 3. YOLOv11 Detection Architecture

YOLOv11 is employed as the core object detection model in the proposed wildlife detection system due to its efficiency and unified detection framework. The model follows a single-stage detection architecture that integrates feature extraction, object localization, and classification into one end-to-end network. This unified approach eliminates the need for separate region proposal stages, thereby simplifying the detection pipeline. Input images captured from camera traps are first fed into the backbone network of YOLOv11.

The backbone consists of multiple convolutional layers that progressively learn hierarchical feature representations from the input data. Early convolutional layers extract low-level features such as edges, contours, and textures, while deeper layers capture high-level semantic features related to animal shape, size, and posture. These learned features provide rich visual representations essential for accurate detection.

The feature maps generated by the backbone network preserve important spatial and contextual information required for object localization. These feature maps are then forwarded to the detection head of the model. The detection head processes the extracted features to predict bounding box coordinates that precisely localize animals within the image. In addition to localization, the detection head estimates class probabilities for each detected object, enabling accurate species identification. All predictions, including bounding boxes and class labels, are generated simultaneously in a single forward pass through the network. This one-stage detection strategy significantly reduces computational complexity and inference time.

Due to its optimized architecture, YOLOv11 achieves high detection speed while maintaining strong accuracy levels. This balance between speed and accuracy makes the model highly suitable for real-time wildlife monitoring applications. The architecture performs reliably under varying illumination conditions, including both daytime visible-light images and nighttime infrared images. The unified and efficient design of YOLOv11 ensures consistent detection performance, enabling continuous wildlife surveillance and supporting effective conservation and monitoring efforts.



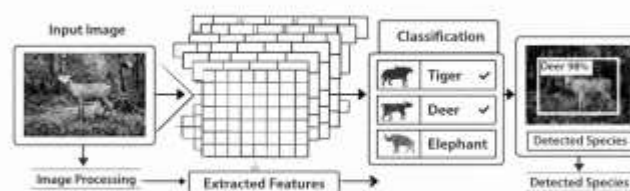
**Fig 5:** System Architecture

#### 4.Feature Extraction and Classification

Feature extraction is automatically performed by the convolutional layers of the YOLOv11 model during the detection process. These layers analyze the input images and learn hierarchical visual representations without manual intervention. Initial convolutional layers capture low-level patterns such as edges and contours, while deeper layers extract more complex visual features. The extracted features represent important characteristics including animal shape, texture, size, and structural patterns. These feature representations help the model distinguish animals from background elements commonly present in forest environments.

The feature maps generated by the network retain both spatial and contextual information required for accurate detection. These maps preserve the relative position of animals within the image while maintaining semantic details. Based on the extracted feature representations, the system performs classification of detected objects. Each detected animal is categorized into a predefined wildlife class learned during the training phase. The classification process is fully automated and operates within the same detection framework.

This automated feature extraction and classification process eliminates the need for manual labeling during inference and reduces continuous human supervision. The integrated approach improves detection consistency across different environmental and lighting conditions. The system performs effectively for both daytime and nighttime images. By combining feature learning and classification within a single unified model, the system enhances detection accuracy and reliability. This design supports efficient and scalable wildlife monitoring and conservation applications.



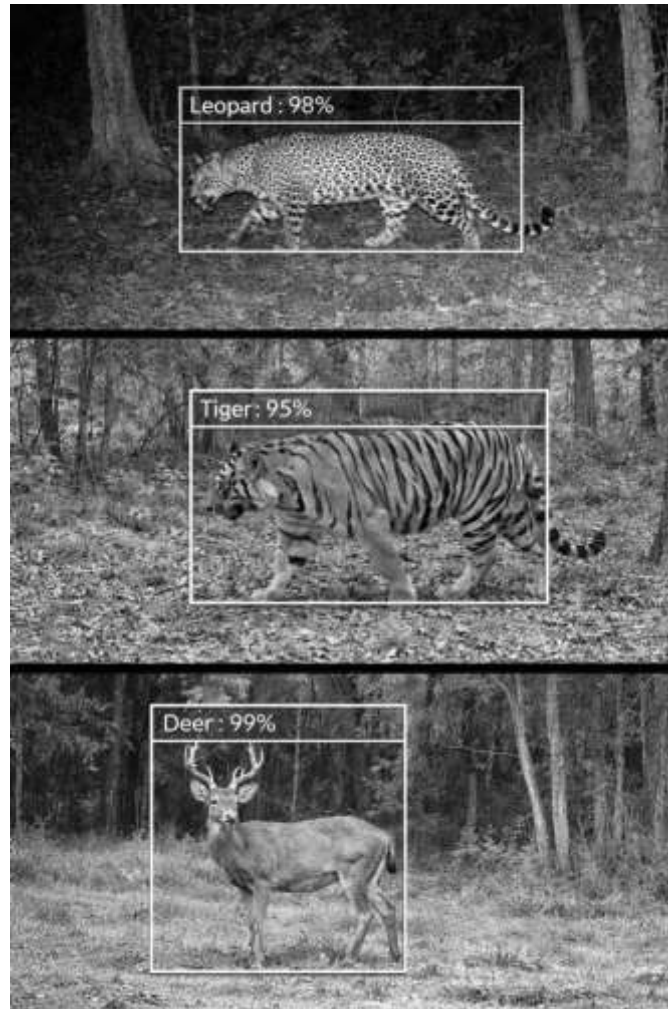
**Fig 6:** Feature Extraction and Classification Process

#### 5.Bounding Box Localization

Bounding box localization is an essential component of the proposed wildlife detection system. It is used to precisely identify and locate animals within the captured camera trap images. Once an animal is detected, the system draws a rectangular bounding box around the object. This bounding box defines the exact spatial position of the animal within the image. Accurate localization helps distinguish animals from complex forest backgrounds and overlapping visual elements.

Along with bounding boxes, the system also displays class labels and confidence scores for each detected animal. The class label indicates the identified wildlife species, while the confidence score represents the detection reliability of the model. These additional details help in confirming the correctness of detection results. By visually marking detected animals, the system enables easy identification of animal presence in both daytime and nighttime images.

Bounding box localization further supports the analysis of animal movement patterns and spatial distribution across forest regions. By observing the location and frequency of detected animals, forest officials and researchers can study wildlife behavior more effectively. Accurate localization improves wildlife tracking and aids in long-term ecological analysis. Overall, this approach enhances monitoring efficiency and supports data-driven wildlife conservation and management efforts.



**Fig 7:** Wildlife Detection with Bounding Boxes

## 6. Output Visualization Module

The output visualization module plays an important role in presenting the detection results in a clear and understandable format. This module displays annotated images in which detected wildlife are highlighted using bounding boxes along with corresponding species labels. The visual annotations allow users to easily recognize detected animals without requiring technical expertise in deep learning models.

The generated outputs are designed to be simple and visually interpretable for forest officials, researchers, and monitoring personnel. Each detected animal is clearly marked, making it easier to verify detection accuracy and species identification. The visualization results help users quickly assess wildlife presence in both daytime and nighttime images captured by camera traps.

By analyzing these annotated outputs, users can monitor animal activity patterns and identify frequently visited regions within forest areas. This information supports effective decision-making for wildlife conservation and forest management. The visualization module therefore enhances situational awareness and provides practical insights for conservation planning and monitoring operations.

## IV. CONCLUSION

In this paper, an automated wildlife detection system based on the YOLOv11 deep learning model has been proposed and successfully implemented. The system is capable of detecting and classifying wildlife from both daytime visible-light images and night-time infrared images obtained through camera trap installations. This dual capability ensures continuous monitoring across different lighting conditions. By automating the processes of wildlife detection and localization, the proposed system significantly reduces the need for manual image analysis. This automation improves monitoring efficiency and minimizes human effort, particularly when handling large volumes of camera trap data. The system effectively identifies wildlife species and accurately localizes them within images. Experimental results demonstrate that the proposed system achieves high detection accuracy and reliable performance under varying environmental and illumination conditions. The findings confirm the robustness and effectiveness of the YOLOv11-based approach for wildlife monitoring applications. Overall, the proposed system provides a practical and efficient solution for wildlife conservation, forest surveillance, and mitigation of human-wildlife conflict. The system is scalable, computationally efficient, and suitable for real-world deployment, making it a valuable tool for modern wildlife monitoring and conservation efforts.

## V. REFERENCES

- [1] **A. Kumar, R. Ramesh, and S. N. Omkar**, “Deep learning-based wildlife detection and classification using camera trap images,” *International Journal of Computer Vision and Image Processing*, vol. 14, no. 1, pp. 1–15, **2024**.
- [2] **P. S. Reddy, K. V. Rao, and M. S. Babu**, “Automated wildlife monitoring system using YOLO-based object detection,” *Journal of Intelligent Systems*, vol. 33, no. 2, pp. 245–258, **2024**.
- [3] **J. Redmon, S. Divvala, R. Girshick, and A. Farhadi**, “You Only Look Once: Unified, Real-Time Object Detection,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, **2016**.
- [4] **M. Norouzzadeh, A. Nguyen, M. Kosmala, et al.**, “Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning,” *Proceedings of the National Academy of Sciences (PNAS)*, vol. 115, no. 25, pp. E5716–E5725, **2018**.
- [5] **S. Beery, G. Van Horn, and P. Perona**, “Recognition in Terra Incognita,” *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 456–473, **2018**.
- [6] **S. Beery, Y. Liu, D. Morris, and A. Piavis**, “Synthetic examples improve generalization for rare classes in camera trap datasets,” *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–10, **2020**.
- [7] **R. Girshick**, “Fast R-CNN,” *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 1440–1448, **2015**.
- [8] **A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao**, “YOLOv4: Optimal speed and accuracy of object detection,” *arXiv preprint arXiv:2004.10934*, **2020**.

### Copyright & License:



Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.