

Zone-Wise Wildlife Detection and Population Analysis Using Computer Vision

M. Vinodhini^{#1}, Dr. P. Anitha^{#2}, Dr. N. Geetha^{#3}

[#]Department of Computer Applications, Vellalar College for Women, Thindal, Erode, Tamil Nadu, India

¹vinodhinimurugasan2005@gmail.com, ²anitha.p@vcw.ac.in, ³geetha.n@vcw.ac.in

Abstract—

Wildlife monitoring is essential for biodiversity conservation, ecological research, habitat management, and protection of endangered species. Traditional wildlife population monitoring methods often require extensive fieldwork, manual observation, and expert analysis, making them time-consuming and difficult to scale across large geographical regions. The development of computer vision, machine learning, and deep learning provides an opportunity to automate wildlife detection and population analysis using camera-trap images and videos. This paper presents a review of computer-vision-based approaches for zone-wise wildlife detection and population analysis. The proposed concept divides a wildlife reserve, forest, or protected area into geographical zones and uses camera traps or other imaging devices to collect data from each zone. Deep-learning models such as YOLO, Faster R-CNN, CNNs, ResNet, EfficientNet, Vision Transformers, and related object-detection techniques can be used to identify animal species and count detected individuals. The resulting information can be combined with the location of each camera to produce zone-wise species distributions, abundance estimates, and population trends. Recent research demonstrates that automated wildlife detection can substantially reduce the manual effort required to process large camera-trap datasets. However, challenges remain in handling small and partially visible animals, different lighting conditions, dense vegetation, similar-looking species, class imbalance, changing habitats, and limited training data. Population estimation also requires consideration of imperfect detection and repeated observations. Future systems can combine computer vision, geographic information systems, edge AI, IoT sensors, and explainable artificial intelligence to provide real-time and zone-specific wildlife monitoring.

Keywords—Wildlife Detection; Computer Vision; Deep Learning; Camera Traps; Animal Detection; Species Classification; Population Estimation; Object Detection; YOLO; Biodiversity Monitoring

I. INTRODUCTION

Wildlife plays an important role in maintaining ecological balance and biodiversity. Monitoring the distribution, abundance, and population trends of animals is therefore essential for conservation planning, habitat management, prevention of poaching, and assessment of environmental changes. However, conventional wildlife monitoring generally depends on field surveys, manual observations, physical tracking, tagging, and expert analysis. These approaches can require considerable time, financial resources, and human effort.

Camera traps have become an important technology for non-invasive wildlife monitoring. They can automatically capture images or videos when animals pass through a particular area. A large camera-trap deployment can generate thousands or millions of images, creating a major data-processing challenge. Recent studies have shown that computer vision can automate several stages of this process, including filtering empty images, detecting animals, identifying species, and counting individuals.

Deep learning has significantly improved automated wildlife recognition. Convolutional Neural Networks (CNNs), object-detection models, transfer learning, and transformer-based architectures can learn visual characteristics from wildlife images. For example, studies have evaluated CNN architectures for wildlife classification, while more recent research has investigated object detection for complex forest scenes and rare species.

An important extension of wildlife detection is **population analysis**. Detecting an animal in an image does not necessarily mean that the animal represents a new individual. The same animal may appear in several images or cameras. Therefore, population analysis requires methods for counting, tracking, estimating abundance, and accounting for repeated observations. Recent research has investigated automated counting, distance estimation, camera-trap distance sampling, occupancy models, and landscape-scale population trends.

The proposed concept of **zone-wise wildlife detection** divides a large wildlife reserve into smaller geographical zones. Each zone can contain one or more camera traps. Images collected from these cameras are processed using computer-vision algorithms to determine the species present and the number of detections. The results can then be associated with individual zones to generate a spatial wildlife database.

For example, a reserve can be divided into Zone A, Zone B, Zone C, and Zone D. If elephant detections are concentrated in Zone A while deer detections are distributed across Zones B and C, the system can produce a zone-wise wildlife distribution map. Over time, repeated observations can be used to identify changes in species occurrence and relative abundance.

Therefore, the integration of computer vision, camera traps, geographical information, and population-analysis techniques can provide an automated framework for large-scale wildlife monitoring.

II. LITERATURE SURVEY

Tabak et al. [1] investigated the accessibility and transferability of machine-learning algorithms for identifying animals in camera-trap images. The researchers used approximately three million camera-trap images from 18 studies across 10 U.S. states. Two deep neural networks were developed: one for recognizing 58 species and another for determining whether an image was empty or contained an animal. The reported accuracies were 96.8% and 97.3%, respectively. The study also emphasized the problem of model transferability between locations.

Inference: A model trained in one zone may perform differently in another zone because of differences in vegetation, camera position, and environmental conditions, so transferability is a key concern for zone-wise monitoring.

Norouzzadeh et al. [2] developed a deep active-learning system for species identification and counting in camera-trap images. The research addressed the problem of manually reviewing millions of images and the difficulty of obtaining large labelled datasets. Active learning was used to reduce the amount of manual annotation required. The study also highlighted the tendency of models to overfit to specific camera locations.

Inference: Active learning can reduce manual annotation effort, and awareness of overfitting to camera locations provides an important foundation for zone-wise systems where cameras may be installed in different habitats.

Schindler and Steinhage [3] investigated identification of animals and recognition of their actions in wildlife videos using deep learning. Camera traps were used to capture animals such as deer, wild boar, foxes, and hares, including nighttime recordings. The research demonstrated that deep learning could be used not only for animal identification but also for recognizing animal actions.

Inference: Deep learning can recognize animal actions as well as identities, which can be extended in zone-wise systems to study species distribution together with behavioral activity.

Yang et al. [4] developed an adaptive deep-learning approach for automatically filtering empty camera-trap images. Their approach focused on reducing the need for powerful computing infrastructure and large manually labelled datasets. Using 29,192 training samples, the reported overall error was 2.69%.

Inference: Empty-image filtering is an important preprocessing stage because wildlife camera traps frequently generate images without animals.

Kutugata et al. [5] investigated automatic camera-trap classification using wildlife-specific deep learning for nilgai management. The study addressed the problem of manually labelling large numbers of camera-trap images. The research demonstrated the applicability of deep learning to species-specific wildlife management.

Inference: Specialized deep-learning models can be developed for particular species or geographical regions.

Stančić et al. [6] evaluated 36 pre-trained CNN models for wildlife classification using Eurasian lynx camera-trap images collected in Croatia. The images varied substantially in quality and the dataset was highly imbalanced.

Inference: Transfer learning is useful for wildlife classification, and suitable pretrained architectures must be selected for camera-trap data.

Johanns et al. [7] proposed AUDIT, an automated approach for estimating the distance between cameras and detected animals. The approach combined monocular depth estimation with an alignment procedure to estimate metric distances without requiring reference images. The reported mean absolute distance-estimation error was 0.9864 m in their evaluation.

Inference: Camera-to-animal distance estimation is particularly useful for population-density estimation because distance is an important factor in several ecological estimation methods.

Zhu et al. [8] investigated class-incremental learning for wildlife biodiversity monitoring. Traditional models are generally trained using a fixed set of species, and adding new species may require retraining the entire model. The proposed incremental-learning approach aimed to recognize new species while reducing catastrophic forgetting of previously learned species.

Inference: Incremental learning is important for long-term zone-wise monitoring because new species may appear in a zone over time.

Norman et al. [9] studied whether CNN-based species classification could generalize across different habitat conditions within camera-trap surveys. The research examined variation in tropical forest disturbance and investigated whether training-data configuration affected model generalization.

Inference: Models must remain reliable when environmental conditions differ between zones, which is a major issue for zone-wise wildlife detection.

Bothmann et al. [10] developed an automated wildlife image-classification pipeline using active learning. Their work showed that upstream object detection can improve classification and that active learning can help use human annotation resources more efficiently. The authors also provided a practical Python package, improving the potential for ecological applications.

Inference: Active learning combined with upstream object detection can improve wildlife image classification while using human annotation resources more efficiently.

Cunha et al. [11] investigated long-tail visual recognition of animal species in camera-trap images. Wildlife datasets often contain many images of common species but relatively few images of rare species. Their proposed approach addressed this imbalance using specialized sampling strategies.

Inference: Rare species can be the most important species to detect, so long-tail recognition methods are particularly relevant to conservation.

The DeepWILD study [12] developed a three-stage deep-learning approach for detecting, classifying, and counting animals in camera-trap videos. The system used object detection, classification, and an individual-counting method. It achieved 96.88% mAP for detection at IoU 0.5 and predicted an exact or ± 1 count for 87% of the evaluated cases when based on detection results.

Inference: Combining species identification with automated counting is directly related to population analysis.

Song et al. [13] benchmarked eight object-detection algorithms for recognizing 15 wild bird species in complex forest scenes using 9,717 camera-trap images. The study compared Faster R-CNN, Cascade R-CNN, RetinaNet, FCOS, RepPoints, ATSS, Deformable-DETR, and Sparse R-CNN. Cascade R-CNN achieved the best performance, with a reported mAP of 0.693, while Cascade R-CNN with a Swin-T backbone achieved 0.704.

Inference: Detecting small and camouflaged animals in complex natural backgrounds remains difficult.

Twining et al. [14] used long-term camera-trapping data to estimate landscape-scale population trends for six species in New York State. Hierarchical models, single-species occupancy models, and Royle-Nichols models were used to account for observation uncertainty. The study produced spatially explicit predictions of occurrence and abundance.

Inference: Camera-trap observations can be converted into meaningful landscape-scale population trends rather than simply counting detections.

Henrich et al. [15] investigated the combination of camera traps, deep learning, and camera-trap distance sampling for large-scale wildlife density estimation in temperate forests. The study used 895,019 manually classified photographs from 10 protected areas in Germany and evaluated the DeepFaune classification system together with camera-trap distance sampling.

Inference: Large-scale automated population-density estimation is a recent development and is particularly relevant to zone-wise wildlife analysis.

The following Table I reviews fifteen representative studies published between 2020 and 2026, highlighting their datasets, techniques, objectives, key findings, and limitations.

TABLE I. COMPARATIVE ANALYSIS OF REPRESENTATIVE STUDIES

Author & Year	Data / Dataset	Main Technique	Main Objective	Key Finding	Limitation / Challenge
Tabak et al. (2020)	3 million images, 18 studies	Deep Neural Networks	Species and empty-image classification	High classification accuracy	Transferability between locations
Norouzzadeh et al. (2021)	Camera-trap images	Deep Active Learning	Species identification and counting	Reduces annotation requirements	Requires labelled examples
Schindler & Steinhage (2021)	Wildlife videos	Deep Learning	Animal/action recognition	Supports animal and behavior recognition	Limited species/video conditions
Yang et al. (2021)	29,192 training samples	Adaptive Deep Learning	Empty-image filtering	2.69% overall error	Training-data requirements
Kutugata et al. (2021)	Nilgai camera traps	Wildlife-specific DL	Automatic species classification	Supports species management	Species-specific application
Stančić et al. (2022)	Eurasian lynx images	36 pretrained CNNs	Wildlife classification	Transfer learning effective	Imbalanced/image-quality variation
Johanns et al. (2022)	Camera-trap images/videos	Depth estimation	Animal-distance estimation	MAE $\approx$ 0.99 m in test setting	Environmental generalization
Zhu et al. (2022)	Camera-trap datasets	Incremental Learning	New-species recognition	Reduces catastrophic forgetting	Computational/memory constraints
Norman et al. (2023)	Tropical forest images	CNN	Habitat-generalized classification	Examined cross-habitat generalization	Habitat variation
Bothmann et al. (2023)	Wildlife camera traps	DL + Active Learning	Automated classification	Reduces manual annotation	Training data still needed
Cunha et al. (2023)	Camera-trap datasets	Long-tail DL	Rare-species recognition	Improved minority-class recognition	Rare training samples
DeepWILD (2023)	Camera-trap videos	Detection + Classification + Counting	Detection and population estimation	96.88% mAP at IoU 0.5	Counting/classification errors
Song et al. (2024)	9,717 bird images	Object Detection	Bird detection	Cascade R-CNN/Swin-T strongest	Small/camouflaged birds
Twining et al. (2024)	Long-term camera traps	Occupancy + abundance models	Population trends	Spatial population trends	Imperfect detection
Henrich et al. (2026)	895,019 images, 10 protected areas	Deep Learning + CTDS	Wildlife density estimation	Large-scale automated estimation	Cross-species/site robustness

III. RESEARCH GAP

Existing research on wildlife monitoring using computer vision mainly focuses on animal detection, species classification, image filtering, and counting from camera-trap images and videos. Although deep learning models such as CNN, YOLO, transfer learning, and transformer-based approaches provide good detection performance, most studies address these tasks separately. Limited research combines wildlife detection with **zone-wise geographical analysis and population-related estimation** in a single system.

Another important gap is the difficulty of handling different habitats, lighting conditions, occlusion, camera positions, and rare species. Repeated detection of the same animal can also

lead to inaccurate population estimates. Many existing systems provide image-level or camera-level counts without considering spatial distribution across different zones and changes over time.

Therefore, there is a need for an integrated computer vision system that can detect wildlife, identify species, count animals, associate detections with specific geographical zones, and analyze population trends over time. Such a system can support automated monitoring, reduce manual effort, and provide useful information for wildlife conservation and habitat management.

IV. CHALLENGING ISSUES

- **Variation in Environmental Conditions:** Changes in lighting, weather, vegetation, and background can reduce the accuracy of wildlife detection.
- **Animal Occlusion:** Animals may be partially hidden by trees, bushes, or other animals, making detection and identification difficult.
- **Species Similarity:** Some wildlife species have similar physical characteristics, which can lead to incorrect species classification.
- **Limited Training Data:** Rare species may have fewer images available for training, resulting in poor recognition performance.
- **Class Imbalance:** Common species may dominate the dataset, while rare species are underrepresented.
- **Accurate Population Estimation:** Repeated images of the same animal can result in over-counting. Therefore, image-based counts may not directly represent the actual population.
- **Zone Identification:** Correctly associating detected animals with geographical zones requires reliable location information from cameras or other sources.
- **Real-Time Processing:** Processing large numbers of camera-trap images and videos requires significant computational resources.
- **Camera Placement:** Different camera angles, distances, and heights can affect detection and classification accuracy.
- **Data Storage and Management:** Large-scale wildlife monitoring generates a huge volume of images and videos, requiring efficient storage, processing, and management.
- **Model Generalization:** A model trained in one habitat may not perform equally well in another habitat with different environmental conditions.
- **Privacy and Data Security:** Wildlife monitoring systems should securely manage location and observation data, particularly when sensitive conservation-area information is involved.

V. MOTIVATION

Wildlife monitoring is important for protecting biodiversity, managing habitats, and understanding changes in animal populations. Traditional wildlife monitoring methods require considerable human effort, time, and resources. Camera traps provide large amounts of wildlife images, but manually examining and counting these images is difficult and time-consuming.

Recent developments in computer vision and deep learning provide an opportunity to automate wildlife detection and species identification. However, many existing systems focus mainly on detecting or classifying animals without providing detailed zone-wise population analysis. Therefore, there is a need for an automated system that can identify wildlife, count detected animals, associate them with specific geographical zones, and analyze population trends.

The motivation of this research is to develop a computer vision-based system that reduces manual monitoring effort and provides useful information about wildlife distribution and population changes across different zones. Such a system can assist researchers, forest departments, and conservation organizations in making data-driven decisions for wildlife protection and habitat management.

VI. SCOPE OF THE STUDY

The scope of this study is to develop a computer vision-based system for automated wildlife detection and population analysis across different geographical zones. The system can process images or videos collected from camera traps and identify wildlife using deep learning techniques.

The study includes **animal detection, species classification, counting, zone-wise distribution analysis, and population trend analysis**. Each detected animal can be associated with a particular zone based on camera or location information. The collected data can then be used to compare wildlife occurrence and population-related trends between different zones and time periods.

The system can be designed to support multiple wildlife species and can be extended as new training data become available. It can also provide visual or statistical reports to assist researchers and wildlife authorities in monitoring animal populations.

The study primarily focuses on automated image-based monitoring and does not claim that simple image counts represent exact animal population size. Accurate population estimation may require additional ecological methods, repeated observations, animal identification, and appropriate statistical models. Future extensions can include real-time monitoring, individual animal recognition, mobile or edge-based processing, and integration with GIS for advanced wildlife distribution analysis.

VII. APPLICATIONS OF ZONE-WISE WILDLIFE DETECTION AND POPULATION ANALYSIS USING COMPUTER VISION

The proposed system can be applied in various areas of wildlife monitoring, conservation, and habitat management:

- **Wildlife Population Monitoring:** Automatically detects and counts animals from camera-trap images to support population-related analysis.
- **Zone-Wise Wildlife Distribution:** Identifies the distribution of different species across forest zones and helps compare wildlife activity between areas.
- **Endangered Species Monitoring:** Supports the identification and monitoring of rare or endangered species in protected habitats.
- **Forest and Habitat Management:** Provides information about animal presence and distribution that can assist forest authorities in habitat management.
- **Wildlife Conservation:** Helps conservation researchers study changes in species occurrence and population trends over time.
- **Human-Wildlife Conflict Management:** Identifies areas with frequent wildlife activity and can support early-warning and conflict-reduction measures.

- **Camera-Trap Data Analysis:** Reduces the time and effort required to manually examine large numbers of wildlife images and videos.
- **Research and Ecological Studies:** Provides structured data for researchers studying animal behavior, habitat usage, species distribution, and seasonal variations.
- **Illegal Activity Monitoring:** Camera-trap systems can support the detection of unusual human activity in protected wildlife areas, subject to applicable monitoring and privacy requirements.
- **Decision Support for Wildlife Authorities:** Generates zone-wise reports and statistics that can assist authorities in planning conservation and monitoring activities.

VIII. CONCLUSION

Computer vision has become an important technology for automated wildlife monitoring. The literature reviewed from 2020 to 2026 demonstrates significant progress in animal detection, species classification, image filtering, counting, distance estimation, incremental learning, and population-density analysis.

Traditional manual analysis of camera-trap data is difficult to scale because large wildlife-monitoring projects can generate millions of images. Deep-learning techniques can reduce this workload by automatically identifying relevant images and detecting animal species. Research has also demonstrated that object detection and automated counting can support population-related analysis.

The proposed zone-wise approach extends these capabilities by connecting each wildlife detection to a specific geographical area. This makes it possible to generate zone-level information about species occurrence, abundance, distribution, and population trends. Such information can assist conservation researchers and wildlife managers in understanding how animals are distributed across a protected area.

However, accurate population analysis requires more than simply counting detected images. Repeated observations, imperfect detection, camera-to-animal distance, environmental variation, rare species, and individual-animal identification must be considered. Recent ecological studies have demonstrated the importance of combining camera-trap observations with statistical population and density-estimation techniques.

Overall, the combination of computer vision, deep learning, camera traps, GIS, and ecological modelling provides a promising foundation for automated zone-wise wildlife monitoring. Such systems can reduce manual effort, improve monitoring efficiency, and provide spatially and temporally organized information for wildlife conservation.

IX. FUTURE RESEARCH DIRECTIONS

Based on the research gaps and challenges identified in this review, the following directions can be explored in future work:

- **Real-Time Zone-Specific Monitoring:** Future systems can process camera-trap data in real time to provide up-to-date, zone-specific information about wildlife occurrence and activity.

- **Integration with GIS:** Combining detections with geographic information systems can support spatial wildlife distribution maps and advanced zone-wise analysis.
- **Edge AI and IoT Sensors:** Mobile or edge-based processing together with IoT sensors can reduce the need to transfer and process very large volumes of images centrally.
- **Individual Animal Recognition:** Recognizing individual animals can reduce repeated counting of the same animal across images and cameras, leading to more reliable population estimates.
- **Statistical Population and Density Estimation:** Combining detections with ecological methods such as camera-trap distance sampling and occupancy models can convert image counts into meaningful abundance and density estimates.
- **Explainable Artificial Intelligence:** Explainable models can help researchers and wildlife authorities understand and trust automated species identification and counting results.
- **Robustness Across Habitats and Rare Species:** Models should be improved to remain reliable across different habitats, lighting conditions, camera positions, and for rare species with limited training data.

REFERENCES

- [1] Tabak, M. A., Norouzzadeh, M. S., Wolfson, D. W., Sweeney, S. J., Vercauteren, K. C., Snow, N. P., Halseth, J. M., Di Salvo, P. A., & others. (2020). Improving the accessibility and transferability of machine learning algorithms for identification of animals in camera trap images: MLWIC2. *Ecology and Evolution*, 10, 10374–10383.
- [2] Norouzzadeh, M. S., Morris, D., Beery, S., Joshi, N., Jojic, N., & Clune, J. (2021). A deep active learning system for species identification and counting in camera trap images. *Methods in Ecology and Evolution*, 12, 150–161.
- [3] Schindler, F., & Steinhage, V. (2021). Identification of animals and recognition of their actions in wildlife videos using deep learning techniques. *Ecological Informatics*, 61, 101215.
- [4] Yang, D.-Q., et al. (2021). An adaptive automatic approach to filtering empty images from camera traps using a deep learning model. *Wildlife Society Bulletin*.
- [5] Kutugata, M., Baumgardt, J., Goolsby, J. A., & Racelis, A. E. (2021). Automatic camera-trap classification using wildlife-specific deep learning in nilgai management. *Journal of Fish and Wildlife Management*, 12(2), 412–421.
- [6] Stančić, A., Vyroubal, V., & Slijepčević, V. (2022). Classification efficiency of pre-trained deep CNN models on camera trap images. *Journal of Imaging*, 8(2), 20.
- [7] Johanns, P., Haucke, T., & Steinhage, V. (2022). Automated distance estimation for wildlife camera trapping. *Ecological Informatics*, 70, 101734.
- [8] Zhu, H., et al. (2022). Class incremental learning for wildlife biodiversity monitoring in camera trap images. *Ecological Informatics*, 68, 101760.
- [9] Norman, A. J., et al. (2023). Can CNN-based species classification generalise across variation in habitat within a camera trap survey? *Methods in Ecology and Evolution*, 14, 242–251.
- [10] Bothmann, L., Wimmer, L., Charrakh, O., Weber, T., Edelhoff, H., Peters, W., Nguyen, H., Benjamin, C., & Menzel, A. (2023). Automated wildlife image classification: An active learning tool for ecological applications. *Ecological Informatics*, 77, 102231.
- [11] Cunha, F., dos Santos, E. M., & Colonna, J. G. (2023). Bag of tricks for long-tail visual recognition of animal species in camera-trap images. *Ecological Informatics*, 76, 102060.

- [12] *DeepWILD: Wildlife identification, localisation and estimation on camera trap videos using deep learning.* (2023). *Ecological Informatics*, 75, 102095.
- [13] Song, Q., Guan, Y., Guo, X., Guo, X., , Y., Wang, H., Ge, J., Wang, T., & Bao, L. (2024). Benchmarking wild bird detection in complex forest scenes. *Ecological Informatics*, 80, 102466.
- [14] Twining, J. P., Kramer, D., Perkins, K. A., & Fuller, A. K. (2024). Landscape-scale population trends in the occurrence and abundance of wildlife populations using long term camera-trapping data. *Biological Conservation*, 290, 110398.
- [15] Henrich, J., et al. (2026). Camera traps and deep learning enable efficient large-scale density estimation of wildlife in temperate forest ecosystems. *Remote Sensing in Ecology and Conservation*, 12(1), 148–163.