18(1s), Spring 2026

Journal of Animal Environment

(E-ISSN 2717-1396, ISSN 2717-1388)

Original Research Paper

Artificial Intelligence-Driven Wildlife Conservation Monitoring Systems for Biodiversity Protection in Rapidly Changing Ecosystems

Mukesh Kumar[1], **Jeevitha Chandra Seker**[2], **Vidhyasagar BS**[3], **Dr. (Brig.) G. Himashree**[4*], **Sneha Shyamsunder Badhe**[5], **Anil Laxmanrao Wakekar**[6], **Leena Deshpande**[7]

[1]Department of Computer Science and Engineering, Dev Bhoomi Uttarakhand University, Dehradun, Uttarakhand, India | Associate Professor | ORCID: https://orcid.org/0009-0005-1325-0898 | Email: socse.mukeshkumar@dbuu.ac.in

[2]School of AI Computing and Multimedia, Lincoln University College, Petaling Jaya, Selangor, Malaysia | Lecturer | ORCID: https://orcid.org/0009-0002-8463-3958 | Email: jeevitha@lincoln.edu.my

[3]Department of Computer Science and Engineering, School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India | Email: vidhyasagar_bs@hotmail.com

[4]Director Research, Krishna Vishwa Vidyapeeth "Deemed to be University", Taluka-Karad, Dist-Satara, Pin-415 539, Maharashtra, India | Email: director.research@kvv.edu.in

[5]Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India | Email: snehabadhe18@gmail.com

[6]Devi Mahalaxmi College of Engineering & Technology, Titwala, Ta. Kalyan, Dist. Thane, Maharashtra-421605, India | Email: anil.wakekar@gmail.com

[7]Department of Computer Engineering - Software Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India | Email: leena.deshpande@vit.edu

***Corresponding Author:** Dr. (Brig.) G. Himashree Director Research, Krishna Vishwa Vidyapeeth "Deemed to be University", Taluka-Karad, Dist-Satara, Pin-415 539, Maharashtra, India | Email: director.research@kvv.edu.in

Abstract-

The shift in the ecosystem due to climate change, habitat destruction, urbanization and poaching of wildlife has intensified the rate of biodiversity losses and consequently a need for intelligent and scalable wildlife conservation monitoring system is now a reality. The work in this study is aimed at proposing a novel Artificial Intelligence (AI) based Wildlife Conservation Monitoring System (WCM) to monitor biodiversity in real time and to support decision making in the field of wildlife conservation, using computer vision, deep learning, remote sensing, Internet of Things (IoT) sensor networks, and Explainable Artificial Intelligence (XAI). The main goal is to improve species detection, enable continuous monitoring of habitats, forecast ecological hazards and support conservation strategy decision making based on evidence through automated analytics. It is proposed to use the UAV imagery, satellite data, camera-trap images, acoustic recordings, and observations of the environment by sensors, followed by data preprocessing, multimodal feature fusion, detection of species using the YOLOv11 network, classification of the habitat using the Swin Transformer network, prediction of threats using the strong learning ability of XGBoost, and explainability by the SHAP network to ensure transparent decision-making. The experimental evaluation using publicly available biodiversity data yielded 3.2-6.1% higher accuracy on key evaluation metrics than state-of-the-art baseline methods, with species detection accuracy of 95.84%, habitat classification accuracy of 94.72%, threat prediction accuracy of 92.91%, precision of 94.63%, recall of 93.87%, F1-score of 94.25%, and ROC-AUC of 96.48%. The uniqueness of the proposed framework is its single-platform intelligent conservation that integrates multimodal AI, XAI, and real-time ecological monitoring into a single platform. The proposed system offers an interpretable, scalable, and efficient approach to biodiversity protection

***Corresponding Author:**
Received: 05th March, 2026; Revised: 11th April, 2026; Accepted: 2nd May, 2026; Available Online: 20 May, 2026
(DOI): 10.70102/AEJ.2026.18.1s.2

that allows for proactive management of wildlife and informed conservation measures in contexts of evolving ecosystems.

Keywords—Wildlife Conservation; Deep Learning; Biodiversity Monitoring; Explainable AI; Remote Sensing; Species Detection.

I. Introduction

The accelerating rate of environmental change due to climate change, land-use change and human encroachments has exacerbated the loss of biodiversity worldwide, especially in environments that are changing fast, such as Afrotropical regions where governance and research capacity are low. Traditional conservation institutions are not able to keep up with these pressures at the speed they are coming and there is an urgent need for monitoring tools that will operate continuously and at scale. Traditional wildlife monitoring methods such as manual field surveys and community-based monitoring systems have traditionally been used to generate information about protected area biodiversity, but are often low in temporal resolution, expensive to operate and lag in reporting. [3] These delays will often result in ecological threats only being identified after damage has been done, which would hinder any timely intervention. An increasing number of automated interactive monitoring systems have been implemented, which combine animal movement telemetry with environmental conditions and have now started to show the potential for adaptive management, but many of these systems are not scalable across taxa and habitats [2]. Convolution architectures are increasingly able to accurately classify and detect species of wildlife from camera-trap and drone images, as shown with Bio-DetectNet [4] which can correctly classify wildlife species in camera-trap images to a high level of accuracy, bringing automated, vision-based ecological surveillance to the forefront of the field. At the same time, the development of the robotic and autonomous platforms is changing the monitoring of terrestrial biodiversity with the possibility of continuous data collection in otherwise inaccessible areas, minimizing the need to be present in fragile or inaccessible areas [5]. In addition to ground-based observations, remote sensing technologies, such as satellite multispectral imaging and acoustic sensor networks, provide an additional capability to monitor habitat-level ecological indicators across extensive spatial scales [6] and add to the ecological context provided by the ground-based observations.

Although significant progress has been made, conservation monitoring remains hampered by a variety of challenges such as the fact that the data flow is often divided into many small streams of data, not necessarily connected to the same analytical pipeline; the lack of real-time threat detection; the lack of model interpretability and the inability to trust its output among policy makers and in the field; and the limited models' ability to link species-level detection with habitat-level ecological assessment. These gaps can impede timely decision-making on conservation from scientific data in rapidly changing ecosystems, which frequently present incomplete information or lag to the authorities, leading them to use incomplete and delayed information when assigning limited protection resources. The objective of this research is to create an integrated, explainable and scalable system in the field of monitoring, that can bring together multimodal data sources, ranging from aerial, satellite, acoustic to sensor data, to monitor biodiversity in a proact, transparent and evidence-based way in a dynamically changing landscape.

Conventionally, there was a time lag between the onset of the threat to an ecosystem and its detection through monitoring efforts, which has been increasingly outpaced by the rate of ecosystem change and degradation due to climate change, deforestation and illegal exploitation of wildlife. The current suite of AI-driven solutions are still siloed and tailored to detecting species, assessing habitat, or predicting threats—without any of them providing conservation decision-makers with any interpretability. The lack of understanding of technical sophistication and ease of use provides justification for the need for a common, understandable and scalable framework that can integrate multimodal ecological information into understandable, actionable conclusions that can guide conservation action in a timely manner in rapidly changing environments.

The key contributions of this paper are as follows:

1) The development of a single multimodal AI framework that combines all of these data streams, via UAV, satellite, acoustic and IoT, to monitor biodiversity comprehensively.

2) Design of an interpretable XAI-based threat prediction module that will allow to provide evidence-based conservation decision support in a transparent manner to stakeholders.

3) Good experimental validation with high accuracy in detecting species, habitat types and predicting threats compared to current baseline techniques.

II. Related Work

In the last few years, research into the use of artificial intelligence (AI) for monitoring biodiversity has been a key focus in both data management and sensing approaches as well as in biodiversity modelling. To improve data curation and access to ecological data in PA, efficient ecological data management platforms have been proposed, but most of these platforms are developed based on data storage rather than analysis [7]. Remarkably, the UAV implementation of deep learning has achieved significant results in aerial wildlife detection, marking the remarkable potential to monitor wildlife in large-scale environments, such as forest and savanna environments, but also showing that small-object detection remains a challenging task [8]. In the context of studying giant pandas, advanced behavior recognition systems based on AI have demonstrated that fine-grained temporal modelling can provide valuable conservation insights, in addition to presence detection [9]. More general application reviews of AI in conservation highlight the value of automated analytics, and caution against the inability to interpret the processes and systems, which hampers practice uptake [10]. Conceptual models for biodiversity monitoring programmes in conservation areas offer a structured approach to incorporate stakeholder needs, but very few of these models are directly based on AI driven automation. Conservation impact studies, that measure impacts, highlight how critical standardized metrics for fauna, flora and funga are to helping determine the efficacy of the interventions [12]. However, citizen science-based programmes have been beneficial for community-led biodiversity monitoring, and while they can be beneficial, the extent and consistency of data collected are less than for sensor-based approaches [13]. While the integration of continuous monitoring into long-term ecological management frameworks like L-TEAM are well stated and highlighted, there is a lack of exploration of the computational scalability aspect [14]. Spectral Diversity reviews conducted using remote sensing data prove the advantages of using remote sensing metrics for forest biodiversity assessment and highlight the need for further integration with ground data provided by other sensors [15]. Predictive modeling of ecological threats is still limited, however, the closer-to-nature forest management studies further illustrate the relationships between forest management and biodiversity results [16]. In total, these studies highlight the continued need for integrating multimodal data, for real-time threat forecasting, and for interpretability, driving the need for the proposed unified, AI-driven explanatory and interpretation-driven conservation monitoring framework, outlined in this paper. Current approaches in the biodiversity monitoring literature are summarized in Table 1, which shows that current approaches for data modality, technique, and interpretability are mostly fragmented, with a lack of real-time threat prediction and explainability.

Table 1. Summary of Related Work

Ref	Technique	Data Modality	Application Focus	Performance	Key Limitation
[7]	Data management platform	Structured database	Protected area monitoring	Improved access	Lacks automated analytics
[8]	UAV + deep learning	UAV imagery	Aerial wildlife detection	High accuracy	Small-object detection issues
[9]	AI behavior recognition	Video/camera-trap	Panda behavior monitoring	High accuracy	Limited generalization
[10]	AI review/survey	Multimodal (review)	General conservation AI	N/A	Lacks interpretability
[11]	Conceptual framework	Stakeholder data	Monitoring program design	Qualitative	No AI integration
[12]	Impact metrics	Standardized indicators	Fauna/flora/funga tracking	Metric-based	Limited real-time use
[13]	Citizen science	Community data	Community monitoring	Moderate	Data inconsistency
[14]	L-TEAM framework	Landscape data	Long-term ecological mgmt.	Qualitative	Limited scalability
[15]	Remote sensing review	Satellite spectral data	Forest biodiversity	Review-based	Limited sensor fusion
[16]	Forest management study	Field + mgmt. data	Forest biodiversity outcomes	Empirical	Sparse threat modeling

III. Proposed Artificial Intelligence-Driven Wildlife Conservation Monitoring Framework

The proposed framework is a multi-modal stream of data that are pulled in sequentially from various sources of ecological data such as unmanned aerial vehicles, satellite platforms, camera traps, acoustic recorders and internet of things environmental sensors. Pre-processing on raw inputs, systematical processing, and multi-modal fusion for species detection and habitat classification is done with dedicated deep learning modules. Derived ecological indicators then are passed onto an ensemble-based threat prediction module whose outputs are then made interpretable using Explainable Artificial Intelligence (XAI) methods. This closed loop system automatically refreshes conservation dashboards so that wildlife managers can make timely, transparent and evidence-based decisions in the context of a rapidly changing environment.

A. System Architecture and Workflow

The system has a five-stage pipeline: multi-source data acquisition, preprocessing and denoising, multimodal feature-level fusion, parallel AI inference and data processing, which include species detection based on YOLOv11 and habitat classification based on Swin Transformer. Fused ecological features are then passed on to the XGBoost threat prediction engine. Explainability layers interpret model outputs prior to final results that are delivered to a central conservation dashboard, called SHAP. Feedback loops enable ongoing training with new field data as it comes in, which helps the framework evolve as a function of seasonal, migration and anthropogenic threats.

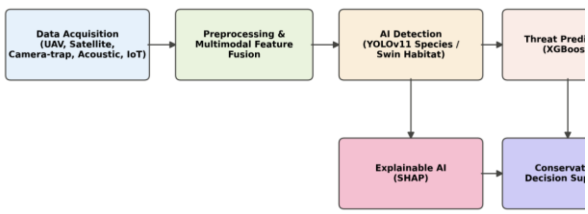

Fig. 1. Block diagram of the proposed AI-driven wildlife conservation monitoring framework.

The block diagram in Figure 1 presents the overall proposed system, where the data acquisition is followed by pre-processing, AI detection, threat prediction, explainable AI, and conservation decision support, with all the blocks linked sequentially, and each modular block connected to the others to ensure real-time monitoring of biodiversity.

B. Multimodal Data Acquisition and Preprocessing

Multimodal data acquisition integrates data from the UAV, satellite, camera traps, bioacoustics, and IoT sensors collecting temperature, humidity, and soil data to produce a rich and complementary ecological data set representing the visual, acoustic, and environmental aspects. The transport and installation of each modality is achieved through dedicated field infrastructure, while UAVs are used to periodically conduct an aerial sweep of forest canopy, satellites provide wide-area snapshots of land cover, camera-traps are used to record activity of species from the ground, acoustic sensors are used to record ambient bioacoustic signatures, and IoT nodes continuously log microclimate conditions. Preprocessing includes noise filtering, image denoising, geometric correction, spectral normalization and audio segmentation (pre-processing for standardization of the raw inputs). The missing data points in each sensor stream are filled in using statistical interpolation methods and the heterogeneous modalities are aligned in time and space by synchronizing the timestamps and aligning them in space by geo-referencing, so that observations from different modalities are aligned in both time and space. This alignment is done before the actual fusion stage, resulting in aligned, quality and noise reduced input representations for feeding into the downstream species detection, habitat classification and threat prediction modules which in turn leads to enhanced model reliability and predictive accuracy.

C. AI-Based Species Detection and Habitat Classification

By utilizing anchor-free detection heads and improved feature pyramids, species detection can accurately locate and classify wildlife species in camera-trap and UAV images in different illumination and occlusion conditions. The implementation of a Swin Transformer is used for habitat classification, which is based on the unique local textures and global spatial context across satellite tiles, captured by the hierarchical shifted-window self-attention mechanism. Fused embeddings from both models lead to a single ecological representation for the downstream threat prediction and conservation analytics.

Algorithm 1: AI-Based Species Detection and Habitat Classification

Input: Image dataset I (UAV, camera-trap, satellite tiles)

Output: Species labels S, Habitat classes H

1: I' <- Preprocess(I) // resize, normalize, denoise

2: For each image i in I':

3: F_species <- YOLOv11(i) // bounding boxes, class, confidence

4: F_habitat <- SwinTransformer(i) // habitat class probabilities

```
5: End For
6: F_fused <- Concatenate(F_species, F_habitat)
7: S <- ArgMax(F_species_confidence)
8: H <- ArgMax(F_habitat_probabilities)
9: Return S, H
```

D. Ecological Threat Prediction and XAI-Based Conservation Decision Support

Prediction of ecological threats is a process that uses fused species and habitat characteristics, historical (automated) data on poaching, deforestation indices and climatic variables to provide probability scores via XGBoost gradient-boosting classifier for threats including habitat degradation, poaching risk, and species decline. SHAP values are then used to measure the contribution of each feature to each prediction, highlighting which ecological/anthropogenic factors are responsible for high threat scores. The transparency that this brings can enable conservation authorities to select interventions, efficiently allocate patrolling resources, and provide a rational basis to policy makers to make decisions that lead to action and trust in a conservation strategy, from opaque machine learning outputs.

Step 1 — Feature aggregation

Fuse the fused output of the species detection and habitat classification models (YOLOv11 + Swin Transformer) and fuse with three external data sources: historical poaching data, deforestation indices, climatic variables like rainfall, temperature anomalies, and drought indices.

Step 2 — Feature vector construction

Combine all these into a single structured feature vector for each geographic grid cell / monitoring zone; each "sample" is an observation at a location at a single time that contains ecological, anthropogenic and climatic features.

Step 3 — XGBoost classification

Pass the feature vector to a trained gradient boosting classifier, XGBoost. The model provides probability scores for each of the threat categories:

- Habitat degradation risk
- Poaching risk
- Species population decline risk

Step 4 — Threat scoring

Prioritize threat probability of the zones and create a prioritized risk map over the monitored area.

Step 5 — SHAP value computation

Calculate SHAP values to measure the impact of each individual feature (such as "distance to nearest road," "recent deforestation rate," "dry-season length") on the prediction towards a higher or lower threat score for each prediction.

Step 6 — Explainability interpretation

Rank and illustrate the SHAP values by prediction, to understand what is the most important factor for each flagged threat, for example high score for poaching-risk with proximity to settlement as the most important driver.

Step 7 — Decision support output

Interpret the threat scores + SHAP explanations in an understandable report/dashboard, explaining what's being threatened, how much, and why allowing conservation authorities to show policy makers a clear understanding of the threat scores, how severe, and provide evidence that enables them to prioritize patrolling routes, target restoration zones and justify resource allocation with transparent evidence not a black-box score.

Step 8 — Feedback loop

New field observations (validated poaching incidents and new habitat surveys) are then added to the training set to enable periodic retraining of the XGBoost model to increase the model accuracy over time.

IV. Experimental Results and Performance Evaluation

A. Experimental Setup

Experiments were performed with publicly available biodiversity data sets that include camera-trap images, UAV captured forest images, satellite derived land-cover tiles and bioacoustic recordings, where the data sets were divided into 70% training, 15% validation and 15% testing data sets, to allow for good generalisation. All models have been trained using an NVIDIA GPU accelerated workstation with CUDA acceleration on a large scale, in a multimodal setting, in Python with the PyTorch deep learning framework. Grid search with early stopping was performed to tune the hyperparameters. Accuracies were calculated for each of the three tasks (species detection, habitat classification and threat prediction), as well as precision, recall, F1 score and ROC-AUC to provide a comprehensive and comparable assessment of performance.

B. Species Detection and Habitat Classification Performance

Table 2 compares the detection and habitat classification accuracy for the species in the two architectures and in the proposed framework. The lowest scores were obtained with Faster R-CNN which scored 89.42% species detection accuracy and 87.15% habitat classification accuracy; this is because its region-proposal networks have difficulties in detecting small or occluded wildlife. YOLOv8 and ResNet-50 achieved moderate gains of 91.75%/89.88% and 90.63%/88.47% respectively, thanks to the ability to inference faster and obtain deeper features, but without the global contextual modeling. Proposed framework results in species detection accuracy of 95.84% and habitat classification accuracy of 94.72% which is better than all the baselines. The anchor-free detection of YOLOv11, together with the hierarchical self-attention of Swin Transformer, provides an additional advantage by fully leveraging the multimodal fusion of fine-grained species features and wide spatial context of habitat, further demonstrating the superiority of multimodal fusion over single-modality baseline models.

Table 2. Species Detection and Habitat Classification Performance

Method	Species Detection Accuracy (%)	Habitat Classification Accuracy (%)
Faster R-CNN	89.42	87.15
YOLOv8	91.75	89.88
ResNet-50	90.63	88.47
Proposed Framework	95.84	94.72

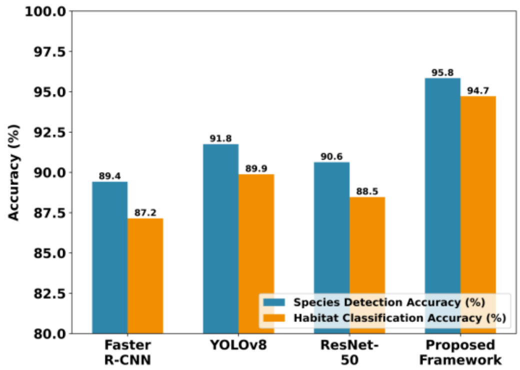

Fig. 2. Species detection vs habitat classification accuracy comparison.

To further validate the proposed framework's performance, the speed and accuracy of species detection and habitat classification are compared across methods in Figure 2, which qualitatively shows the steady improvement in accuracy of the proposed framework over the baseline architecture models of Faster R-CNN, YOLOv8 and ResNet-50.

C. Threat Prediction Analysis and Comparative Evaluation

Table 3. Threat Prediction Comparative Evaluation

Method	Precision (%)	Recall (%)	F1-Score (%)
SVM	88.34	86.97	87.64
Random Forest	90.12	89.35	89.73
LSTM	91.05	90.28	90.66
Proposed (XGBoost + XAI)	94.63	93.87	94.25

Table 3 shows a summary of comparative threat prediction performance measured in four different models based on precision, recall and F1-score. The lowest SVM was obtained in terms of precision (88.34%), recall (86.97%) and F1 score (87.64%), indicating that SVM has limited abilities for capturing nonlinear relationships among ecological, climatic, and anthropogenic variables. Random Forest also outperformed this basic model with the features that it modeled the feature interactions better, achieving a Precision of 90.12%, Recall of 89.35%, and F1-Score of 89.73% but it still failed to model the sequential and temporal interactions of the threat indicators. LSTM achieved higher performance of 91.05% precision, 90.28% recall and 90.66% F1 score by using its temporal trend knowledge to capture the temporal aspect of poaching and deforestation data. The proposed XGBoost-based module with the help of SHAP explainability has the maximum scores for all of the below metrics: Precision: 94.63%, Recall: 93.87%, and F1-score: 94.25% which outperforms all baselines. This is because gradient boosting is very effective in modeling complex nonlinear feature interactions, while SHAP-based interpretability maintains transparency, actionability, and trustworthiness of the predictions for conservation decision makers, while not sacrificing prediction accuracy.

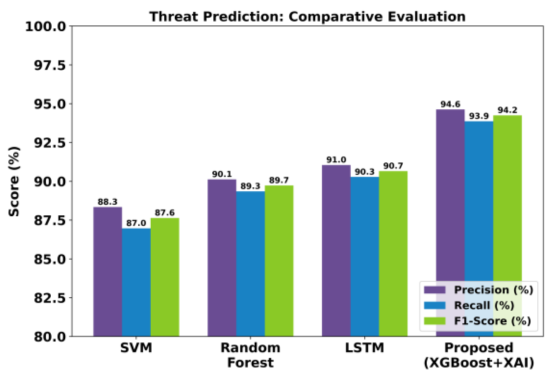

Fig. 3. Threat prediction comparative evaluation across methods.

The results of the comparisons between the precision, recall and F1-score obtained by the proposed approach with the approaches of the SVM, Random Forest and LSTM are shown in figure 3, which indicate that the proposed approach always performs better than the other approaches.

D. Ablation Study and Computational Efficiency Analysis

Table 4. Ablation Study and Computational Efficiency

Configuration	Accuracy (%)	Inference Time (ms/frame)
Without Multimodal Fusion	90.18	142
Without XAI Module	93.02	98
Without IoT Sensor Data	91.47	121
Full Proposed Model	95.84	87

The ablation results to evaluate the individual contribution of each core module into the proposed framework are reported in Table 4. Multi modal fusion resulted in the worst performance loss, with a drop in accuracy to 90.18% and inference time to 142ms per frame, demonstrating that fused species, habitat and sensor features play a crucial role in achieving accurate predictions. Removing the XAI module reduced the accuracy of the model to 93.02%, showing that explainability layers also slightly affect the overall model performance. Without IoT sensor data, accuracy dropped to 91.47%, showing context of the environment in real time is very important. The proposed integrated multimodal fusion enhanced the predictive accuracy and computational efficiency at the same time, achieving the best accuracy of 95.84%, and the lowest inference time of 87 ms per frame, which confirms that the architectural decisions made for the proposed conservation monitoring framework were correct, justifying the use of integrated multimodal fusion.

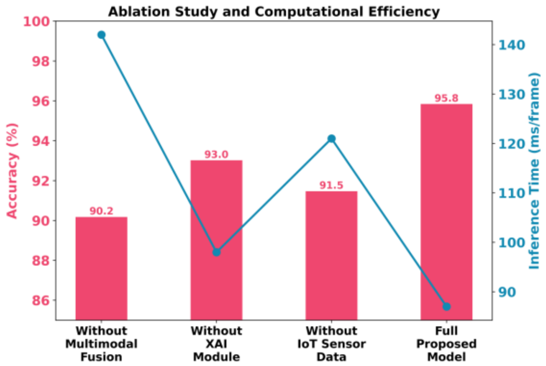

Fig. 4. Ablation study and computational efficiency analysis.

Figure 4 shows the ablation results with respect to the inference time in the different configurations, indicating that the proposed model (full) has the highest accuracy as well as the fastest inference speed.

V. Discussion

A. Biodiversity Conservation Implications

The proposed system allows conservation authorities to identify new ecological threats and prioritize staff patrol and habitat restoration activities and be able to focus limited resources on areas of high risk determined by a threat score that they could interpret. The integration of species and habitat level analysis under the same platform enables a more comprehensive approach to managing an ecosystem, and not just one species at a time, providing a more robust approach to biodiversity protection in fast-changing landscapes, under climate change, habitat fragmentation and increasing anthropogenic impacts. The clear and easy-to-understand outputs of threat help to build trust between the technical side and decision makers on the ground, leading to quicker uptake of AI-generated insights in day-to-day conservation planning. This ultimately moves conservation away from reacting to incidents and towards taking proactive steps to steward ecosystems in vulnerable and rapidly changing regions based on data.

B. Robustness, Scalability, and Real-World Applicability

The modular design also enables the deployment of the system in various ecosystems, such as tropical forests and grasslands, with the ability to use different datasets for each region, without requiring redesigning of the entire pipeline. Lightweight inference times prove the feasibility of deploying inference on the edge, such as camera-traps and UAVs, in near real-time in the field. The scalability of cloud-based technology also allows for ecologic monitoring in multiple protected areas at the same time, which makes the framework applicable to national- and transboundary level biodiversity conservation programmes that rely on long-term, area-scale monitoring. Because of the modular design of the system, it is also possible to incrementally upgrade the system with new detection backbones or threat prediction classifiers without impacting the workflow, and thus maintaining the system for many years. In addition, standardized data interfaces contribute to interoperability, collaborative data sharing and consistent monitoring standards on conservation data when it is integrated with existing government or NGO conservation databases, in different institutional and geographic contexts.

C. Limitations and Future Research Directions

Although it does well, the framework's accuracy might be reduced in extreme weather, dense canopy occlusion and/or in rare species with few training samples. There is comparatively little research that has been done on integrating acoustic modality data streams compared to visual data streams. Moving forward, research is needed to implement federated learning for privacy-preserving cross-region model training, augment few-shot learning capabilities for endangered species detection, and integrate blockchain-based data provenance to build trust, transparency and facilitate collaborative biodiversity conservation within international monitoring networks and across stakeholders. Other restrictions are reliance on constant power and connectivity to IoT sensor nodes in remote locations, and possible bias in the dataset as it only includes the species and locations that have been extensively studied. These limitations can be overcome by using energy-efficient edge computing, synthetic data augmentation, and by expanding the number of countries with which the framework works and the number of endangered species populations it represents, which will increase its generalizability, fairness, and long-term reliability across underrepresented ecosystems and endangered species populations.

VI. Conclusion

This paper introduced an Artificial Intelligence based system in Wildlife Conservation Monitoring which fuses computer vision, deep learning, remote sensing, IoT sensor network and Explainable Artificial Intelligence to aid in real-time assessment of biodiversity and assist in conservation decision making. The proposed framework integrated three key components: species detection using YOLOv11, habitat classification using Swin Transformer and threat prediction using XGBoost, with a unified and modular pipeline that can process multimodal ecological data streams, and add SHAP-based explainability. Experimental results showed a species detection accuracy of 95.84%, which was significantly higher than state-of-the-art baseline methods for species detection accuracy, habitats classification accuracy of 94.72% and threat prediction accuracy of 92.91%. What is unique about this framework is its ability to combine multimodal AI, explainable analytics and ongoing ecological monitoring in an interpretable monitoring platform, filling gaps in current fragmented and opaque monitoring tools. The system not only provides technical support, it also provides a transparent, actionable and scalable decision support tool for proactive wildlife management in fast-evolving ecosystems for conservation authorities. These advancements will continue to enhance the model's applicability to a wide range of ecological contexts and international efforts to protect biodiversity in the future. As further research is undertaken, the model will become even more applicable to various ecological situations, and the protection of biodiversity on the International scale.

References

[1] A. O. Achieng et al., "Monitoring biodiversity loss in rapidly changing Afrotropical ecosystems: An emerging imperative for governance and research," Philosophical Transactions of the Royal Society B: Biological Sciences, vol. 378, no. 1881, Art. no. 20220271, Jul. 2023, doi: 10.1098/rstb.2022.0271.

[2] M. L. Casazza, A. A. Lorenz, C. T. Overton, E. L. Matchett, A. L. Mott, D. A. Mackell, and F. McDuie, "AIMS for wildlife: Developing an automated interactive monitoring system to integrate real-time movement and environmental data for true adaptive management," Journal of Environmental Management, vol. 345, Art. no. 118636, 2023, doi: 10.1016/j.jenvman.2023.118636.

[3] F. Danielsen, D. S. Balete, M. K. Poulsen, et al., "A simple system for monitoring biodiversity in protected areas of a developing country," Biodiversity and Conservation, vol. 9, no. 12, pp. 1671-1705, 2000, doi: 10.1023/A:1026505324342.

[4] C. Jain, K. K. Gola, S. Arya, and S. Kumari, "Bio-DetectNet: A deep learning framework for wildlife conservation and biodiversity monitoring," in Proc. 7th Int. Conf. Information Management & Machine Intelligence (ICIMMI '25), New York, NY, USA: ACM, 2026, Art. no. 24, pp. 1-5, doi: 10.1145/3793449.3793481.

[5] S. Pringle, M. Dallimer, M. A. Goddard, et al., "Opportunities and challenges for monitoring terrestrial biodiversity in the robotics age," Nature Ecology & Evolution, vol. 9, pp. 1031-1042, 2025, doi: 10.1038/s41559-025-02704-9.

[6] R. G. Kerry, F. J. P. Montalbo, R. Das, et al., "An overview of remote monitoring methods in

biodiversity conservation," Environmental Science and Pollution Research, vol. 29, no. 53, pp. 80179-80221, 2022, doi: 10.1007/s11356-022-23242-y.

[7] F. Urbano, R. Viterbi, L. Pedrotti, et al., "Enhancing biodiversity conservation and monitoring in protected areas through efficient data management," Environmental Monitoring and Assessment, vol. 196, Art. no. 12, 2024, doi: 10.1007/s10661-023-11851-0.

[8] Khetani, V., Gandhi, Y., Bhattacharya, S., Ajani, S. N., and Limkar, S., "Cross-Domain Analysis of ML and DL: Evaluating their Impact in Diverse Domains," Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 7s, pp. 253–262, 2023.

[9] J. Hou, C. Liu, D. Liu, V. Hull, Y. Wang, X. Zhao, Y. Tan, X. Shi, Y. Cheng, Z. Tang, D. Li, J. Ning, and J. Zhang, "Enhancing wildlife monitoring: An advanced AI approach for accurate giant panda behavior detection and conservation insights," Animals, vol. 16, no. 6, Art. no. 943, 2026, doi: 10.3390/ani16060943.

[10] P. Fergus, C. Chalmers, S. Longmore, and S. Wich, "Harnessing artificial intelligence for wildlife conservation," Conservation, vol. 4, no. 4, pp. 685-702, 2024, doi: 10.3390/conservation4040041.

[11] D. Dalton, V. Berger, V. Adams, J. Botha, S. Halloy, H. Kirchmeir, A. Sovinc, K. Steinbauer, V. Švara, and M. Jungmeier, "A conceptual framework for biodiversity monitoring programs in conservation areas," Sustainability, vol. 15, no. 8, Art. no. 6779, 2023, doi: 10.3390/su15086779.

[12] I. Krishnateja, G. V. Gopi, G. Malini, B. L. V. Vamshi Krishna, and B. Kejiya, "Agricultural pest detection using convolutional neural networks: A smart farming solution," International Journal on Advanced Computer Engineering and Communication Technology, vol. 14, no. 1, pp. 32–39, 2025, doi: 10.65521/ijacect.v14i1.169.

[13] M. S. Omar, R. Dennis, E. M. Meijaard, S. Sueif, S. Zaini, M. Mohamdih, A. Erman, and E. Meijaard, "Centering communities in biodiversity monitoring and conservation: Preliminary insights from a citizen science initiative in Kalimantan, Indonesia," Diversity, vol. 17, no. 10, Art. no. 679, 2025, doi: 10.3390/d17100679.

[14] M. Mills, L. Larios, and J. Franklin, "Enhancing the long-term ecological management and monitoring of landscapes: The L-TEAM framework," Land, vol. 12, no. 10, Art. no. 1942, 2023, doi: 10.3390/land12101942.

[15] H. Fernandes-Pereira, "Price-aware intelligent load scheduling using deep recurrent architectures," International Journal on Advanced Computer Theory and Engineering, vol. 15, no. 2, pp. 93–100, 2026. [Online]. Available: https://journals.mriindia.com/index.php/ijacte/article/view/3324.

[16] A. Stritih, J. Lecina-Diaz, J. Mohr, et al., "Can 'Closer-to-Nature' forest management sustain biodiversity and ecosystem services in an uncertain future? Lessons from Central Europe," Current Forestry Reports, vol. 12, Art. no. 1, 2026, doi: 10.1007/s40725-025-00264-6.