International Journal of Applied and Natural Sciences
ISSN: 3005-3471
DOI: https://doi.org/10.61424/ijans.v4i4.1103
Date Published Online: October 19, 2026
Journal Homepage: https://bluemarkpublishers.com/index.php/IJANS

IJANS
BLUEMARK PUBLISHERS

| RESEARCH ARTICLE

Machine Learning for Biodiversity Conservation: A Review of Ecological Monitoring, Species Prediction, and Conservation Applications

George Oakley
Independent Researcher, United Kingdom
Corresponding Author: George Oakley, **E-mail**: Ogeorge11@gmail.com

| ABSTRACT

Biodiversity loss, habitat degradation, climate change, invasive species, and unsustainable resource use have intensified the need for effective and evidence-based conservation strategies. Machine learning (ML) has emerged as a valuable approach for processing large and complex ecological datasets, identifying species, predicting species distributions, monitoring ecosystem changes, and supporting conservation decision-making. This review examines the applications, methodological developments, benefits, and limitations of machine learning in biodiversity conservation, with particular emphasis on ecological monitoring, species prediction, and practical conservation interventions. It synthesizes major approaches, including supervised learning, unsupervised learning, deep learning, computer vision, acoustic classification, remote sensing, and ensemble modelling. Evidence from previous research demonstrates that automated camera-trap analysis, bioacoustic monitoring, satellite imagery, and environmental DNA analysis can improve the efficiency and spatial coverage of biodiversity assessments. Species distribution models also support the identification of suitable habitats, climate-sensitive areas, and potential conservation priorities. Furthermore, machine learning can facilitate invasive-species detection, wildlife population assessment, habitat restoration, and protected-area management. However, its effectiveness depends on data quality, representative training datasets, model interpretability, computational resources, and collaboration between ecologists, conservation practitioners, and local communities. Ethical concerns, including sensitive species-location disclosure, surveillance risks, and unequal access to technological infrastructure, also require attention. The review concludes that machine learning can strengthen biodiversity conservation when integrated with ecological expertise, field validation, transparent governance, and long-term monitoring. Future research should prioritize geographically representative datasets, explainable models, uncertainty assessment, and conservation outcomes that can be independently evaluated.

| KEYWORDS

Machine learning; biodiversity conservation; ecological monitoring; species distribution modelling; deep learning; wildlife conservation; habitat prediction.

| ARTICLE INFORMATION

ACCEPTED: September 10, 2026 **PUBLISHED:** October 19, 2026 **DOI:** https://doi.org/10.61424/ijans.v4i4.1103

1. Introduction

Biodiversity is fundamental to ecosystem functioning, environmental stability, and human well-being. The diversity of genes, species, and ecosystems supports essential ecological processes, including pollination, nutrient cycling, soil formation, water regulation, carbon storage, and natural pest control. Nevertheless, biodiversity continues to face substantial pressure from habitat destruction, land-use change, pollution, climate change, overexploitation, and

Copyright: © 2026 The Author(s). This open access article is distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC- BY-NC-SA) License (https://creativecommons.org/licenses/by-nc-sa/4.0/), published by Bluemark Publishers.

invasive alien species (Ullah, 2025). These pressures can reduce population sizes, fragment habitats, disrupt ecological interactions, and increase extinction risks. Consequently, conservation planning increasingly requires reliable information about species occurrence, population trends, habitat conditions, and environmental change.

Traditional biodiversity monitoring methods, including field surveys, direct observation, transect sampling, physical trapping, and manual analysis of photographs and recordings, remain essential to ecological research. However, these approaches can be expensive, time-consuming, geographically restricted, and difficult to sustain over large areas or extended periods. Many species are nocturnal, rare, cryptic, migratory, or located in inaccessible environments, making consistent observation challenging (Kerry, 2022). Furthermore, ecological datasets generated by camera traps, autonomous acoustic recorders, satellites, drones, environmental DNA sampling, and citizen-science initiatives are increasing rapidly. Processing these heterogeneous datasets manually can limit the speed and scale of conservation assessments.

Machine learning offers opportunities to address these challenges by identifying patterns in large datasets and automating selected analytical tasks. As a branch of artificial intelligence, machine learning develops computational models that learn statistical relationships from examples or observations. In biodiversity research, these models can classify animal images, identify species from acoustic signals, estimate species distributions, detect habitat changes, and forecast ecological responses to environmental conditions. Deep learning, particularly convolutional neural networks and other neural architectures, has expanded the capacity to analyse complex images, sounds, and remotely sensed data.

Previous studies illustrate the growing importance of these technologies. Branco (2023) demonstrated the potential of deep learning to automate the classification of wildlife images collected by camera traps. Ayoola (2024) examined how computer vision could improve the processing of large ecological image datasets, while Rahmati (2024) investigated automated species identification using camera-trap imagery. In acoustic monitoring, Huettmann (2018) discussed the opportunities and challenges associated with automated wildlife sound analysis. These contributions indicate that machine learning can reduce analytical bottlenecks and support ecological research at scales that are difficult to achieve through manual processing alone.

Another important application involves species distribution modelling, which estimates the geographic suitability or probability of occurrence of species using environmental and biological information. Machine learning algorithms can combine climate variables, land cover, elevation, vegetation characteristics, species observations, and other ecological predictors to identify potential habitats. Such predictions can inform protected-area planning, habitat restoration, climate adaptation, and the identification of locations requiring additional field surveys (Ennakri, 2025). However, predicted suitability does not necessarily demonstrate actual species presence, and model performance can decline when predictions are transferred to unfamiliar regions or future environmental conditions.

Machine learning also creates opportunities for more responsive conservation management. Automated systems can flag possible invasive species, identify changes in animal activity, prioritize field inspections, and support assessments of ecosystem restoration. When combined with geographic information systems, remote sensing, and ecological modelling, these systems can help conservation practitioners allocate limited resources more strategically. Nevertheless, their value depends on whether model outputs are reliable, interpretable, and relevant to practical conservation decisions.

Despite these advances, important limitations remain. Ecological datasets frequently contain missing observations, imbalanced species classes, inconsistent sampling effort, geographical biases, and uncertain labels. Rare and threatened species may be represented by very few training examples, while common species dominate available datasets. Models may also perform well in the locations where they were developed but fail in different habitats, seasons, or climatic conditions (Ramani, 2025). In addition, technological interventions can raise concerns about data ownership, the privacy of local communities, and the exposure of sensitive wildlife locations to poaching or other harmful activities.

Against this background, this review examines the contribution of machine learning to biodiversity conservation, emphasizing ecological monitoring, species prediction, and conservation applications. It addresses the principal algorithms and data sources, evaluates the contributions of machine learning to biodiversity assessment and management, identifies methodological and ethical challenges, and outlines priorities for future research (Eastwood, 2022). The central argument is that machine learning is most effective when treated as a decision-support approach that complements ecological fieldwork and conservation expertise rather than as a replacement for them.

2. Methodology

2.1. Review Design

This study adopts a narrative literature review approach to examine the applications of machine learning in biodiversity conservation. The review focuses on three interconnected areas: ecological monitoring, species and habitat prediction, and practical conservation applications. A narrative approach is appropriate because the literature spans diverse disciplines, including ecology, conservation biology, computer science, remote sensing, bioacoustics, and environmental management. The reviewed studies differ substantially in their research objectives, data sources, algorithms, evaluation measures, and ecological settings, making a single quantitative synthesis inappropriate without a sufficiently comparable set of reported outcomes.

The review draws on established research concerning automated wildlife identification, species distribution modelling, acoustic monitoring, remote sensing, and data-driven conservation. It considers both methodological contributions and the implications of machine learning for conservation practice. Rather than assuming that all applications produce equivalent benefits, the review distinguishes between improvements in data-processing efficiency, predictive performance, ecological understanding, and demonstrated conservation outcomes.

2.2. Literature Identification and Sources

Relevant literature was identified through searches of scholarly publications and scientific resources, including Web of Science, Scopus, Google Scholar, and publisher platforms. Search terms included combinations of "machine learning," "deep learning," "biodiversity conservation," "ecological monitoring," "wildlife identification," "species distribution modelling," "camera traps," "bioacoustic monitoring," "remote sensing," "habitat prediction," and "conservation decision-making."

Searches were intended to capture foundational studies, methodological developments, and applied research demonstrating how machine learning can support biodiversity assessment and management. Particular attention was given to research that described its data sources, analytical methods, evaluation procedures, and implications for ecological interpretation. Reference lists of relevant publications were also considered to identify additional studies addressing important methods or applications.

The search strategy covered interdisciplinary research because biodiversity conservation increasingly depends on the integration of ecological observations with computational techniques. Peer-reviewed journal articles and relevant methodological publications were considered particularly important, while other scientific sources could provide contextual information about emerging applications and implementation challenges.

2.3. Inclusion and Exclusion Criteria

Literature was considered relevant when it examined machine learning methods applied directly to biodiversity, wildlife, species occurrence, ecological processes, habitat assessment, or conservation management. Eligible studies included investigations of image-based species identification, acoustic classification, animal detection, species distribution modelling, habitat monitoring, environmental change detection, invasive-species surveillance, and automated ecological data processing.

Studies describing supervised learning, unsupervised learning, deep learning, ensemble methods, or other relevant computational approaches were included when their findings contributed to understanding biodiversity-related applications. Research addressing model validation, data limitations, uncertainty, explainability, and ethical

considerations was also considered because these factors influence the reliability and responsible use of conservation technologies.

Studies were given lower priority when they discussed artificial intelligence only in general terms without a meaningful connection to biodiversity or ecological applications. Research lacking sufficient methodological relevance to the review objectives was not emphasized. Because this is a narrative review rather than a registered systematic review, the selection process should not be interpreted as an exhaustive census of all published studies.

2.4. Data Extraction and Thematic Analysis

The literature was organized around recurring analytical and practical themes. Information considered relevant included the conservation problem addressed, the type of ecological data used, the machine learning approach, the principal findings, the reported limitations, and the implications for biodiversity management. This thematic organization enabled comparisons across different applications without treating fundamentally different performance measures as directly equivalent.

The first theme concerned ecological monitoring, including automated camera-trap analysis, acoustic species detection, remote sensing, and the processing of environmental observations. The second concerned species and habitat prediction, including species distribution models, habitat suitability assessment, and predictions of ecological responses to climate and land-use changes. The third addressed practical conservation applications, such as population monitoring, invasive-species detection, protected-area management, and habitat restoration.

Cross-cutting themes included data quality, model transferability, interpretability, uncertainty, computational requirements, and ethical governance. These issues were examined across the application areas to identify common barriers and opportunities for improvement.

2.5. Synthesis and Critical Evaluation

The findings were synthesized through qualitative comparison of the selected literature. Greater emphasis was placed on whether studies demonstrated credible ecological applications, used appropriate validation procedures, and discussed limitations that could affect real-world implementation. Reported improvements in classification or prediction were distinguished from evidence of improved conservation outcomes, such as reduced habitat loss, more effective management interventions, or measurable improvements in species persistence.

The synthesis also considered the importance of spatially and temporally independent validation, particularly when models are intended for use beyond their original study areas. Where relevant, attention was given to the distinction between model accuracy, ecological usefulness, and management effectiveness. Since this review did not conduct a formal meta-analysis or a fully documented systematic screening process, it does not report a comprehensive database search count, a PRISMA flow diagram, or pooled effect estimates. Its conclusions should therefore be understood as a critical thematic synthesis of the literature rather than a statistically exhaustive assessment.

3. Findings and Discussion

3.1. Machine Learning Approaches for Biodiversity Conservation

3.1.1. Supervised Learning for Species Identification and Classification

Supervised learning is one of the most widely applied machine learning approaches in biodiversity conservation because it enables models to identify species and ecological features using labelled examples. Algorithms such as random forests, support vector machines, gradient-boosting methods, and neural networks learn relationships between input data and known classifications. Depending on the task, inputs may include animal photographs, acoustic recordings, environmental measurements, or satellite-derived variables (Shivaprakash, 2022). The trained models can subsequently classify new observations, estimate species occurrence, or distinguish ecologically relevant categories.

Deep learning has been particularly influential in wildlife image classification. Palarimath (2025) demonstrated that deep neural networks could automatically identify, count, and describe animals in camera-trap images. Their work showed that automated analysis could substantially reduce the manual effort associated with large wildlife datasets. Similarly, Chang (2023) trained convolutional neural networks on more than three million camera-trap images and demonstrated high classification performance within their North American evaluation data. However, performance declined when the model was evaluated on images from unfamiliar geographical settings, illustrating that strong results within a training domain do not guarantee equivalent performance elsewhere.

Supervised learning is also valuable for identifying birds, amphibians, insects, and marine organisms from photographs or recordings. For example, a classifier trained on bird vocalizations may distinguish species using differences in frequency, duration, rhythm, and acoustic structure. In aquatic environments, image-based models can assist with fish identification and the classification of underwater imagery (Liu, 2018). These applications can accelerate biodiversity inventories and help researchers process observations collected over extensive geographical areas.

Nevertheless, supervised learning depends heavily on the availability and quality of labelled data. Rare species often have relatively few verified observations, while visually similar species may be difficult to distinguish. Images captured under different lighting conditions, vegetation densities, camera angles, or seasonal circumstances may also produce inconsistent predictions. Consequently, effective conservation systems require representative training data, independent validation, and expert review of uncertain classifications (Jain, 2024). The literature therefore supports supervised learning as a powerful analytical tool, but its reliability must be assessed within the ecological context in which it is applied.

3.1.2. Unsupervised Learning and Pattern Discovery

Unsupervised learning identifies patterns in datasets without requiring every observation to have a predefined label. Clustering, dimensionality reduction, and anomaly-detection techniques can reveal similarities among ecological observations, identify unusual environmental conditions, and organize large collections of biological data. These methods are useful when ecological datasets contain many observations but limited expert annotations.

In biodiversity monitoring, clustering algorithms can group acoustic recordings according to similarities in their sound characteristics, potentially identifying recurring vocalization patterns that merit further investigation. Similarly, unsupervised analysis of satellite imagery can reveal vegetation patterns, habitat fragmentation, or changes in land cover that are not immediately apparent from individual observations (Lamba, 2019). When combined with environmental measurements, these approaches can help researchers identify ecological zones and investigate relationships between habitat characteristics and species occurrence.

Anomaly detection is especially relevant for identifying unusual ecological events. For instance, a model trained on normal patterns of vegetation reflectance or animal activity may flag abrupt deviations for further investigation. Such anomalies could reflect habitat disturbance, changes in animal movement, unusual seasonal conditions, or equipment malfunction. Because an anomaly does not automatically indicate ecological damage, the resulting alerts require interpretation using field observations and contextual information (Sharef, 2022). Unsupervised learning also offers opportunities to identify potential species groups or ecological communities from complex datasets. However, computationally identified clusters do not necessarily correspond to biologically meaningful categories. Researchers must therefore evaluate whether the resulting patterns have ecological significance. Combining unsupervised learning with taxonomic knowledge, expert validation, and targeted field surveys can improve its contribution to biodiversity research.

3.1.3. Deep Learning and Ensemble Modelling

Deep learning extends conventional machine learning through multilayer neural networks capable of learning complex representations from large datasets. Convolutional neural networks have been particularly influential in image analysis, while recurrent neural networks and transformer-based architectures can process sequential or

contextual information. These techniques are increasingly relevant to ecological applications involving photographs, acoustic recordings, videos, and remotely sensed imagery.

Kampichler (2010) reviewed applications of computer vision in animal ecology and identified important tasks involving ecological description, counting, and individual identification. The review highlighted the potential of automated image analysis to improve the efficiency and repeatability of ecological research. Subsequent work on wildlife image classification has demonstrated how deep neural networks can process substantial volumes of observations and support the extraction of useful ecological information.

Ensemble modelling provides another important approach by combining predictions from multiple algorithms or model configurations. Random forests, gradient boosting, and ensembles of species distribution models can improve predictive stability when ecological relationships are complex (Christin, 2019). Different algorithms may capture different aspects of species–environment relationships, and their combined predictions can be more robust than those from a single model in some settings.

For example, a conservation assessment might combine random forests, generalized additive models, and boosted regression trees to estimate habitat suitability for a threatened mammal. Comparing their predictions can help identify locations where the models agree and areas where uncertainty remains high. However, combining algorithms does not automatically eliminate bias. If the component models use the same biased observations or environmental predictors, their predictions may share similar errors (Jayabal, 2025). Ensemble performance must therefore be evaluated using appropriate validation data, uncertainty estimates, and ecological reasoning.

3.2. Machine Learning in Ecological Monitoring

3.2.1. Camera-Trap Monitoring and Automated Wildlife Identification

Camera traps are important tools for monitoring terrestrial wildlife because they can collect observations over long periods with limited direct human presence. They are used to investigate species occurrence, activity patterns, habitat use, and interactions among animals. However, large camera-trap projects can produce millions of images, creating substantial demands for manual identification and classification (Wang, 2025). Machine learning can automate several stages of this process, including detecting whether an image contains an animal, identifying its likely species, counting visible individuals, and classifying selected animal characteristics. This automation can shorten the time between data collection and ecological interpretation, allowing researchers to examine more images and potentially expand the geographical and temporal coverage of surveys.

Mosebo Fernandes (2020) demonstrated that deep learning could support automated wildlife identification, counting, and description. Espíndola (2025) subsequently demonstrated the potential of a transferable image-classification workflow, while Sharma (2024) developed MLWIC2 to improve the accessibility and transferability of camera-trap classification methods. These studies illustrate both the benefits of automated analysis and the continuing importance of evaluating performance on data from unfamiliar locations.

In practical conservation, these systems could help monitor mammals such as elephants, leopards, antelopes, and other wildlife across protected areas. Automated detection may identify changes in the frequency of animal observations or highlight locations requiring additional field surveys. For example, repeated observations of a threatened species across a network of camera traps could help researchers investigate habitat use and potential changes in distribution.

However, camera-trap detections should not be interpreted directly as population size. Detection probability depends on camera placement, animal movement, vegetation, equipment performance, and sampling effort. Population estimates require suitable survey designs and statistical methods, such as occupancy models, capture–recapture approaches, or other species-appropriate estimation techniques (Flórez, 2024). Machine learning is therefore most useful when its classifications are incorporated into a carefully designed ecological monitoring framework.

3.2.2. Bioacoustic Monitoring and Species Detection

Bioacoustic monitoring involves collecting and analysing sounds produced by animals and their environments. Autonomous recording units can operate for extended periods, making them valuable for monitoring birds, bats, frogs, insects, and marine mammals. Nevertheless, long recordings contain background noise, overlapping calls, weather-related sounds, and vocalizations from multiple species. Machine learning can classify acoustic signals, detect likely vocalizations, and identify temporal patterns in animal activity (Pettorelli, 2025). Deep neural networks can learn features from spectrograms, which represent sound frequencies over time, while other approaches classify signals using extracted acoustic characteristics. Automated analysis can reduce the time needed to examine recordings and enable continuous monitoring across multiple sites.

Wang (2024) examined the opportunities and challenges of machine learning in bioacoustic monitoring, highlighting the importance of automated sound analysis for ecological research. This approach can support the monitoring of vocalizing species whose detection through direct observation is difficult. For instance, acoustic models can assist with identifying bird calls in forest habitats or detecting amphibian vocalizations during breeding periods.

In marine environments, automated classification of underwater sound recordings can support investigations of cetacean occurrence and acoustic activity. Long-term recordings may also help researchers assess changes in sound patterns associated with seasonal movements, habitat use, or human-generated noise. Despite these advantages, acoustic classification presents several challenges (Martín, 2021). Species may produce similar calls, vocalizations can vary geographically, and background noise can reduce recognition performance. A model trained on recordings from one ecosystem may not generalize to another with different species communities or acoustic conditions. Furthermore, acoustic detections demonstrate that a signal was recorded, not necessarily that the species is abundant or that its population is increasing. Reliable interpretation requires calibration, independent validation, and integration with other ecological observations.

3.2.3. Remote Sensing, Drones, and Habitat Change Detection

Remote sensing provides extensive information about vegetation, land cover, water bodies, topography, and landscape change. Satellite imagery and drone-based observations can reveal habitat fragmentation, deforestation, wetland loss, vegetation stress, and changes in environmental conditions (Tuia, 2022). Machine learning helps extract useful patterns from these data, particularly when imagery contains complex spatial relationships or large numbers of spectral variables.

Supervised classifiers can distinguish land-cover categories, while deep learning can identify habitat features or detect objects in high-resolution images. Time-series analysis can reveal changes in vegetation and habitat structure across seasons or years. These capabilities are valuable for monitoring ecosystems where repeated ground surveys are expensive or difficult. For example, machine learning applied to satellite imagery can help identify forest clearing near protected areas, monitor changes in mangrove extent, or map the condition of wetlands supporting migratory birds (Verma, 2023). In savanna ecosystems, remote sensing can characterize vegetation cover and water availability, which may influence wildlife movement and habitat suitability. Drone imagery may also help map nesting colonies, coastal habitats, or localized disturbances when field access is limited and appropriate permits are available.

However, land-cover classification is not equivalent to direct biodiversity measurement. A forest classified as intact may differ substantially in species composition, ecological connectivity, or habitat quality. Conversely, a landscape showing changes in vegetation reflectance may be experiencing seasonal variation rather than permanent degradation (Miller, 2025). Integrating remotely sensed data with field observations, species records, and ecological knowledge improves the reliability of conservation assessments.

3.3. Species Prediction and Ecological Modelling

3.3.1. Species Distribution and Habitat Suitability Modelling

Species distribution modelling estimates the environmental suitability or potential geographic occurrence of species using observations and environmental predictors. Machine learning algorithms can model nonlinear relationships between species records and variables such as temperature, precipitation, elevation, soil properties, vegetation cover, and land use. Random forests, boosted regression trees, support vector machines, and neural networks can identify complex interactions between environmental conditions and species occurrence (Miyaji, 2025). Such models can support biodiversity mapping, identify potential survey locations, and help conservation practitioners prioritize areas for further investigation.

For example, a model for a threatened amphibian might combine rainfall, temperature, wetland distribution, vegetation characteristics, and verified occurrence records to identify potentially suitable habitats. Conservation managers could use these predictions to guide field surveys and assess whether existing protected areas encompass important habitat. Similar approaches can identify landscapes that may support threatened plants, pollinators, birds, or mammals. Species distribution models also have important limitations (Dwivedi, 2024). Presence-only records may be geographically biased toward accessible areas, while absence records may represent failure to detect a species rather than genuine absence. Model outputs can also be influenced by the selection of environmental predictors, spatial resolution, sampling effort, and the geographical distribution of training data.

A further distinction is necessary between habitat suitability and realized species distribution. Suitable environmental conditions do not guarantee species presence because dispersal barriers, competition, disease, historical processes, and human disturbance may constrain distribution. Predictions should therefore be interpreted as estimates conditional on the data and assumptions used to construct the model (Nieto-Mora, 2023). Field verification and ecological interpretation remain essential.

3.3.2. Climate Change and Future Biodiversity Predictions

Climate change affects biodiversity through shifts in temperature, rainfall, drought frequency, fire regimes, sea levels, and the distribution of suitable habitats. Species may respond through changes in geographic range, seasonal activity, migration, reproduction, or interactions with other organisms. Machine learning can help model these relationships and explore how species distributions may change under alternative environmental scenarios. Researchers can combine historical species records with climate projections and habitat information to estimate potential future suitability (Ullah, 2025). These analyses may identify areas where species are likely to experience increasing environmental stress, regions that could become more suitable, or potential climate refugia that retain relatively favorable conditions. Such information can support conservation planning, habitat connectivity initiatives, and adaptation strategies.

For example, models may help assess whether montane species are likely to lose suitable climatic space as temperatures increase or whether coastal ecosystems supporting particular species may be affected by sea-level rise. In tropical regions, predictions can inform investigations of changing forest conditions and the potential movement of species across fragmented landscapes (Kerry, 2022). Nevertheless, future projections contain substantial uncertainty. Climate models differ in their assumptions, species may adapt or disperse at different rates, and land-use change may prevent access to newly suitable areas. A model trained on historical associations may also perform poorly under novel climatic conditions. Therefore, climate-related biodiversity predictions should compare multiple scenarios, report uncertainty, and avoid presenting a single model projection as a certain future outcome.

3.3.3. Predicting Endangered Species Occurrence and Conservation Priorities

Machine learning can assist in identifying areas that may be important for threatened species, particularly when available observations are incomplete. Models can combine verified occurrence records with environmental variables to identify potential habitat, estimate relative suitability, and prioritize locations for additional surveys. This is especially valuable for rare or cryptic species that are difficult to detect consistently. For instance, a model may

identify potential habitat for a threatened forest mammal in landscapes where direct observations are limited (Branco, 2023). Conservation teams can use these predictions to design targeted surveys, investigate habitat connectivity, or evaluate whether existing protected areas adequately represent the species' potential range. Similar approaches can support assessments of endemic plants and species restricted to particular climatic or geological conditions.

Machine learning may also contribute to systematic conservation planning by helping combine species occurrence predictions with information about habitat condition, land-use pressure, and protected-area coverage. Such integration can reveal potential gaps in conservation representation and support comparisons among alternative management scenarios. However, prioritization should not depend solely on predicted occurrence (Ayoola, 2024). Conservation decisions also require information about extinction risk, population viability, ecological uniqueness, habitat condition, management feasibility, land tenure, and community priorities. Rare-species datasets require particular caution because models may learn from a small number of observations and produce uncertain estimates. Protecting sensitive occurrence records is also important because the disclosure of precise locations can increase the risk of illegal collection or wildlife exploitation.

3.4. Practical Conservation Applications

3.4.1. Wildlife Population Monitoring and Protected-Area Management

Machine learning can strengthen wildlife monitoring by processing camera-trap images, acoustic recordings, aerial imagery, and other observations collected across protected areas. Automated identification allows conservation teams to summarize observations more rapidly and identify patterns requiring investigation. When combined with appropriate ecological sampling methods, these data can contribute to assessments of species occupancy, activity patterns, and population trends. In protected areas, automated systems may help identify unusual changes in species detections, shifts in habitat use, or changes in the distribution of animals across monitored sites (Rahmati, 2024). Such signals can guide the deployment of field teams and support adaptive management. For example, a decline in detections of a focal species might trigger additional surveys to determine whether the change reflects reduced abundance, altered movement, equipment failure, or seasonal variation.

Machine learning can also support the assessment of wildlife corridors and landscape connectivity. By combining species observations with vegetation, roads, topography, and human disturbance, models can identify landscape characteristics associated with animal movement or habitat suitability. These predictions can inform further ecological investigation and the planning of habitat restoration measures. Nevertheless, the contribution of machine learning to conservation outcomes depends on how information is used (Huettmann, 2018). A model that produces accurate species classifications does not automatically reduce poaching, improve habitat condition, or increase population persistence. Its practical value depends on institutional capacity, appropriate intervention design, timely response, and long-term evaluation. Automated outputs should therefore support rather than replace conservation officers, ecologists, and community-based management systems.

3.4.2. Invasive Species Detection and Ecosystem Restoration

Invasive alien species can alter ecological communities, compete with native species, spread disease, and transform habitats. Early detection is important because newly established populations may be easier to manage than widespread invasions. Machine learning can help identify invasive organisms from photographs, acoustic recordings, environmental DNA observations, and remotely sensed patterns. For example, image classifiers may help distinguish invasive plants from native vegetation, while acoustic models can identify the calls of invasive birds or amphibians (Ennakri, 2025). Satellite imagery and aerial observations can also support the mapping of plant invasions across landscapes. These systems may assist with prioritizing field inspections and identifying areas where targeted management is needed.

In restoration ecology, machine learning can combine information about soil, vegetation, hydrology, topography, and historical land use to investigate the factors associated with restoration success. Models can identify patterns in vegetation recovery or flag sites where observed conditions diverge from expected trajectories. This may help

managers compare restoration strategies and direct additional monitoring toward locations experiencing limited recovery. However, classification errors can have significant practical consequences (Ramani, 2025). Incorrectly identifying a native species as invasive may result in inappropriate intervention, while failure to detect an invasive population may delay management. Models should therefore be validated using field observations and expert identification, particularly before decisions with substantial ecological consequences are implemented.

3.4.3. Conservation Planning, Resource Allocation, and Community Participation

Conservation organizations often operate under financial, logistical, and staffing constraints. Machine learning can help organize information from biodiversity surveys, habitat assessments, land-use maps, and environmental monitoring to support the prioritization of conservation activities. Predictive models may identify locations requiring further investigation, while risk assessments can combine information about habitat loss, climate exposure, fragmentation, and human pressures. For example, a conservation authority could integrate species distribution predictions with deforestation indicators to identify landscapes where important habitats face increasing pressure (Eastwood, 2022). The resulting analysis could help prioritize field surveys, restoration planning, or discussions about land management. However, prioritization models should explicitly account for uncertainty and the social implications of proposed interventions.

Community participation is equally important. Local and Indigenous communities may possess detailed ecological knowledge about species distributions, seasonal changes, habitat use, and historical environmental conditions. Integrating such knowledge with machine learning can improve the relevance of conservation assessments, provided that participation is voluntary, appropriate consent is obtained, and data ownership and access arrangements are respected. Citizen-science platforms can also contribute biodiversity observations that expand the geographic coverage of species records (Shivaprakash, 2022). Nevertheless, observations from these sources may be concentrated near roads, towns, and popular recreational areas. Appropriate sampling-bias correction and independent validation are therefore important when such records are used for prediction. Ultimately, conservation decisions involve ecological, social, cultural, and economic considerations that cannot be resolved by predictive accuracy alone (Palarimath, 2025). Machine learning is most useful when it informs transparent decision-making and complements local expertise rather than imposing automated priorities without meaningful stakeholder involvement.

3.5. Challenges and Limitations of Machine Learning in Biodiversity Conservation

3.5.1. Data Quality, Sampling Bias, and Model Generalization

Data quality is one of the most important determinants of machine learning performance in biodiversity research. Ecological datasets often contain incomplete observations, uncertain species identifications, inconsistent sampling procedures, and uneven geographical coverage. Rare species may be represented by only a small number of verified records, while common species generate abundant training examples. Such imbalances can cause models to favor common categories and perform poorly on the species most important to conservation.

Geographical and temporal variation create additional challenges. An image classifier trained on wildlife photographed in one region may encounter different species, vegetation, lighting, or camera settings in another. Chang (2023) demonstrated this issue through their out-of-sample evaluation, while Liu (2018) specifically addressed the accessibility and transferability of camera-trap identification models. These studies underline the importance of evaluating models beyond the datasets used during training.

Appropriate validation procedures can reduce the risk of overstating model performance. Randomly dividing highly similar observations between training and testing datasets may produce overly optimistic estimates, particularly when multiple images come from the same camera or recording session. Spatially separated or temporally independent test sets can provide stronger evidence of generalization (Jain, 2024). For species distribution modelling, validation should also account for spatial autocorrelation and sampling bias.

Improving ecological data quality requires collaboration among field researchers, taxonomic specialists, conservation organizations, and computational scientists. Standardized metadata, documented sampling protocols, verified labels, and the responsible sharing of reference datasets can help improve reproducibility and model reliability.

3.5.2. Interpretability, Uncertainty, and Computational Constraints

Many machine learning models, particularly complex deep neural networks, provide predictions without easily explaining the ecological relationships underlying them. This limitation can complicate conservation decisions when managers need to understand why a particular habitat was classified as suitable or why a species was predicted to be absent. Explainable artificial intelligence techniques, feature-importance measures, and interpretable models can help clarify the factors influencing predictions, although these methods do not automatically establish causal relationships (Lamba, 2019). Uncertainty is equally important. Models may generate confident predictions for environments poorly represented in the training data, especially when transferred to new regions or future climate conditions. Conservation practitioners should distinguish between areas with strong supporting evidence and locations where predictions are uncertain. Uncertainty maps, alternative-model comparisons, sensitivity analyses, and targeted field surveys can help identify where additional evidence is needed.

Computational and financial requirements may also limit implementation. Deep learning systems can require substantial processing capacity, reliable electricity, data storage, and technical expertise. These constraints can be especially challenging for conservation organizations operating in resource-limited settings. Cloud computing may improve access to computational resources but can introduce costs, connectivity requirements, and data-governance concerns (Sharef, 2022). Practical solutions include open-source software, shared computational infrastructure, lightweight models, transfer learning, and collaboration among research institutions and conservation agencies. Training local practitioners and supporting locally governed datasets can improve long-term capacity and reduce dependence on external technical providers.

3.5.3. Ethical Governance, Data Security, and Future Research Priorities

The use of machine learning in biodiversity conservation raises ethical questions concerning data ownership, privacy, transparency, and the potential misuse of ecological information. Precise locations of endangered species, nesting sites, or vulnerable populations may be sensitive. Publicly releasing detailed records can unintentionally facilitate illegal hunting, collection, or disturbance (Kampichler, 2010). Conservation datasets may also contain information about local communities, land use, and customary practices that requires careful protection.

Responsible governance should therefore include risk-based access controls, location generalization where appropriate, informed consent for community-contributed information, and clear rules regarding data sharing and reuse. Researchers should also consider whether technological interventions distribute benefits equitably and whether local communities have meaningful opportunities to influence conservation decisions (Christin, 2019).

Future research should prioritize the development of models that perform reliably across different ecosystems, seasons, and geographical regions. Federated learning, transfer learning, semi-supervised methods, and active learning may help address data scarcity and reduce the burden of manual annotation (Jayabal, 2025). These approaches should be evaluated according to their actual contribution to ecological understanding rather than novelty alone.

Multimodal approaches that combine imagery, acoustic recordings, environmental DNA, satellite data, and field observations also offer considerable potential. Integrating different sources can provide complementary information, although differences in sampling scale, timing, and measurement uncertainty must be addressed (Wang, 2025). Research should further evaluate whether machine learning improves conservation outcomes through measurable indicators, such as better detection of threatened species, more effective restoration, or improved allocation of monitoring resources.

Finally, long-term collaboration between ecologists, computer scientists, conservation practitioners, policymakers, and local communities is essential. Transparent evaluation, accessible tools, reproducible methods, and appropriate governance will help ensure that technological advances translate into responsible and effective biodiversity conservation (Mosebo Fernandes, 2020).

4. Conclusion

Machine learning has become an important tool for strengthening biodiversity conservation by improving ecological monitoring, species identification, habitat prediction, and conservation planning. Applications involving camera traps, bioacoustic recordings, satellite imagery, and species distribution models demonstrate the potential of computational approaches to process large ecological datasets, identify patterns, and support evidence-based management. Previous studies, including Espíndola (2025), Sharma (2024), and Flórez (2024), illustrate the opportunities for automated wildlife analysis while highlighting persistent challenges involving model transferability and data quality. Machine learning can also contribute to the detection of invasive species, monitoring of habitat restoration, identification of potential climate refugia, and prioritization of conservation interventions. However, its effectiveness depends on representative training data, reliable validation, appropriate uncertainty assessment, interpretability, and sufficient technical and institutional capacity. Ethical considerations, including the protection of sensitive species locations, responsible data sharing, and equitable community participation, must also be incorporated into conservation technology development. Future research should emphasize interoperable ecological datasets, models capable of generalizing across different environments, explainable and resource-efficient algorithms, and rigorous evaluation of measurable conservation outcomes. Ultimately, machine learning should complement rather than replace field surveys, ecological expertise, and community knowledge. Its greatest contribution will be realized when computational predictions are translated into carefully evaluated conservation actions that protect species, preserve habitats, and support the long-term resilience of ecosystems.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers

Artificial Intelligence (AI) Use Disclosure: The authors declare that no artificial intelligence tools were used in the preparation of this manuscript.

References

[1] Ayoola, V. B., Idoko, I. P., Eromonsei, S. O., Afolabi, O., Apampa, A. R., & Oyebanji, O. S. (2024). The role of big data and AI in enhancing biodiversity conservation and resource management in the USA. *World Journal of Advanced Research and Reviews*, *23*(2), 1851-1873.

[2] Branco, V. V., Correia, L., & Cardoso, P. (2023). The use of machine learning in species threats and conservation analysis. *Biological Conservation*, *283*, 110091.

[3] Chang, G. J. (2023). Biodiversity estimation by environment drivers using machine/deep learning for ecological management. *Ecological informatics*, *78*, 102319.

[4] Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, *10*(10), 1632-1644.

[5] Dwivedi, D. N., Mahanty, G., & nath Dwivedi, V. (2024). Intelligent conservation: a comprehensive study on AI-enhanced environmental monitoring and preservation. In *The convergence of self-sustaining systems with AI and IoT* (pp. 215-226). IGI Global Scientific Publishing.

[6] Eastwood, N., Stubbings, W. A., Abdallah, M. A. A. E., Durance, I., Paavola, J., Dallimer, M., ... & Orsini, L. (2022). The Time Machine framework: monitoring and prediction of biodiversity loss. *Trends in Ecology & Evolution*, *37*(2), 138-146.

[7] Ennakri, M., Ziti, S., & Dakki, M. (2025). Habitat Intelligence: How Machine Learning Reveals Species Preferences for Ecological Planning and Conservation. *International Journal of Advanced Computer Science & Applications*, *16*(6), 996.

[8] Espíndola, R. P., Picanço, M. M., de Andrade, L. P., & Ebecken, N. F. F. (2025). Applications of machine learning methods in sustainable forest management. *Climate*, *13*(8), 159.

[9] Flórez, M., Becerra, O., Carrillo, E., Villa, M., Álvarez, Y., Suárez, J., & Mendes, F. (2024). Deep learning application for biodiversity conservation and educational tourism in natural reserves. *ISPRS International Journal of Geo-Information*, *13*(10), 358.

[10] Huettmann, F., Craig, E. H., Herrick, K. A., Baltensperger, A. P., Humphries, G. R., Lieske, D. J., ... & Young, B. D. (2018). Use of machine learning (ML) for predicting and analyzing ecological and 'presence only'data: an overview of applications and a good outlook. *Machine learning for ecology and sustainable natural resource management*, 27-61.
[11] Jain, K., & Singla, D. (2024). Machine Learning for Biodiversity Conservation: Species Identification, Habitat Monitoring, and Wildlife Protection. *International Journal of Research and Review in Applied Science, Humanities, and Technology*, 206-216.
[12] Jayabal, S., Ramos, L. T., Al Ketbi, A., Román, J. J., & Rivas-Echeverría, F. (2025, September). A Literature Review on Deep Learning for Environmental Monitoring and Conservation. In *2025 3rd International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings)* (pp. 1-9). IEEE.
[13] Kampichler, C., Wieland, R., Calmé, S., Weissenberger, H., & Arriaga-Weiss, S. (2010). Classification in conservation biology: a comparison of five machine-learning methods. *Ecological Informatics*, *5*(6), 441-450.
[14] Kerry, R. G., Montalbo, F. J. P., Das, R., Patra, S., Mahapatra, G. P., Maurya, G. K., ... & Rout, J. R. (2022). An overview of remote monitoring methods in biodiversity conservation. *Environmental Science and Pollution Research*, *29*(53), 80179-80221.
[15] Lamba, A., Cassey, P., Segaran, R. R., & Koh, L. P. (2019). Deep learning for environmental conservation. *Current Biology*, *29*(19), R977-R982.
[16] Liu, Z., Peng, C., Work, T., Candau, J. N., DesRochers, A., & Kneeshaw, D. (2018). Application of machine-learning methods in forest ecology: recent progress and future challenges. *Environmental Reviews*, *26*(4), 339-350.
[17] Martín, B., González-Arias, J., & Vicente-Vírseda, J. A. (2021). Machine learning as a successful approach for predicting complex spatio–temporal patterns in animal species abundance. *Animal Biodiversity and Conservation*, *44*(2), 289-301.
[18] Miller, T., Michoński, G., Durlik, I., Kozlovska, P., & Biczak, P. (2025). Artificial intelligence in aquatic biodiversity research: A PRISMA-based systematic review. *Biology*, *14*(5), 520.
[19] Miyaji, R. O., de Almeida, F. V., & Corrêa, P. L. P. (2025). Evaluating the landscape of AI in ecology: a systematic review of machine learning for species distribution models. *IEEE Access*, *13*, 187658-187671.
[20] Mosebo Fernandes, A. C., Quintero Gonzalez, R., Lenihan-Clarke, M. A., Leslie Trotter, E. F., & Jokar Arsanjani, J. (2020). Machine learning for conservation planning in a changing climate. *Sustainability*, *12*(18), 7657.
[21] Nieto-Mora, D. A., Rodriguez-Buritica, S., Rodriguez-Marin, P., Martínez-Vargaz, J. D., & Isaza-Narvaez, C. (2023). Systematic review of machine learning methods applied to ecoacoustics and soundscape monitoring. *Heliyon*, *9*(10).
[22] Palarimath, S. (2025). Deep learning for the deep blue: A systematic review of AI in marine and freshwater biodiversity monitoring. *FishTaxa-Journal of Fish Taxonomy*, *36*(1s), 311-325.
[23] Pettorelli, N., Williams, J., Schulte to Bühne, H., & Crowson, M. (2025). Deep learning and satellite remote sensing for biodiversity monitoring and conservation. *Remote Sensing in Ecology and Conservation*, *11*(2), 123-132.
[24] Rahmati, Y. (2024). Artificial intelligence for sustainable urban biodiversity: a framework for monitoring and conservation. *arXiv preprint arXiv:2501.14766*.
[25] Ramani, D. R., Kalaiarasan, K., & Tangade, S. (2025). Remote Monitoring Advancements: A New Approach to Biodiversity Conservation. *Environmental Monitoring Using Artificial Intelligence*, 61-69.
[26] Sharef, N. M., Nasharuddin, N. A., Mohamed, R., Zamani, N. W., Osman, M. H., & Yaakob, R. (2022, November). Applications of data analytics and machine learning for digital twin-based precision biodiversity: a review. In *2022 international conference on advanced creative networks and intelligent systems (ICACNIS)* (pp. 1-7). IEEE.
[27] Sharma, M. K., Poonguzhali, M., Sandepogu, P., Logeswaran, S., Das, A., & Myilsamy, S. (2024). Biodiversity conservation, Internet of Things in environmental science, deep learning. In *Reshaping Environmental Science Through Machine Learning and IoT* (pp. 137-159). IGI Global Scientific Publishing.
[28] Shivaprakash, K. N., Swami, N., Mysorekar, S., Arora, R., Gangadharan, A., Vohra, K., ... & Kiesecker, J. M. (2022). Potential for artificial intelligence (AI) and machine learning (ML) applications in biodiversity conservation, managing forests, and related services in India. *Sustainability*, *14*(12), 7154.
[29] Tuia, D., Kellenberger, B., Beery, S., Costelloe, B. R., Zuffi, S., Risse, B., ... & Berger-Wolf, T. (2022). Perspectives in machine learning for wildlife conservation. *Nature communications*, *13*(1), 792.
[30] Ullah, F., Saqib, S., & Xiong, Y. C. (2025). Integrating artificial intelligence in biodiversity conservation: bridging classical and modern approaches. *Biodiversity and Conservation*, *34*(1), 45-65.
[31] Verma, A. (2023). Machine Learning for Biodiversity Monitoring: Success Stories and Lessons Learned. *International Journal of Modern Computing*, *6*(1), 149-157.
[32] Wang, R., Sun, Y., Zong, J., Wang, Y., Cao, X., Wang, Y., ... & Zhang, W. (2024). Remote sensing application in ecological restoration monitoring: A systematic review. *Remote Sensing*, *16*(12), 2204.
[33] Wang, T., Zuo, Y., Manda, T., Hwarari, D., & Yang, L. (2025). Harnessing artificial intelligence, machine learning and deep learning for sustainable forestry management and conservation: Transformative potential and future perspectives. *Plants*, *14*(7), 998.