International Journal of Research and Review in Applied Science, Humanities, and Technology
Vol 1, Issue 2 December 2024 ISSN: 3048-975X
https://ijrasht.com/

Machine Learning for Biodiversity Conservation: Species Identification, Habitat Monitoring, and Wildlife Protection

Kritika Jain, Department of Computer Science and Engineering (AI & ML), Panipat Institute of Engineering and Technology (PIET), Samalkha, Panipat, Haryana, India 2822302.cse@pietgroup.co.in

Mansi, Department of Computer Science and Engineering (AI & ML), Panipat Institute of Engineering and Technology (PIET), Samalkha, Panipat, Haryana, India 2822351.cse@pietgroup.co.in

Dhairya, Department of Computer Science and Engineering (AI & ML), Panipat Institute of Engineering and Technology (PIET), Samalkha, Panipat, Haryana, India 2822362.cse@pietgroup.co.in

Bhumika, Department of Computer Science and Engineering (AI & ML), Panipat Institute of Engineering and Technology (PIET), Samalkha, Panipat, Haryana, India 2822370.cse@pietgroup.co.in

Deepak Singla, Department of Computer Science and Engineering (AI & ML), Panipat Institute of Engineering and Technology (PIET), Samalkha, Panipat, Haryana, India divya.cse-aiml@piet.co.in

Abstract—Biodiversity monitoring traditionally depends on labour-intensive fieldwork, including manual review of camera trap imagery, acoustic recordings, and field survey data, creating a persistent bottleneck between the volume of ecological data that can be collected and the capacity to analyse it in support of timely conservation decisions. Machine learning has increasingly been applied to close this gap, automating species identification from images and audio, monitoring habitat change from satellite imagery, and supporting population assessment and anti-poaching efforts. This paper surveys machine learning applications in biodiversity conservation, organising methods into a taxonomy spanning camera trap image classification, bioacoustic species identification, remote sensing-based habitat monitoring, individual animal re-identification, species distribution modelling, and environmental DNA analysis. We review applications, including large-scale camera trap networks, anti-poaching systems, deforestation monitoring, marine biodiversity assessment, and citizen science platforms; summarise benchmarks and datasets used in this domain; and present a case study illustrating an integrated camera trap and anti-poaching alert pipeline. We discuss challenges, including geographic and taxonomic bias in training data, the long-tailed distribution of species observations, and the difficulty of generalising models across ecosystems, and outline future directions connecting conservation machine learning with multimodal monitoring systems and community-based conservation partnerships.

Index Terms—biodiversity conservation, species identification, camera traps, bioacoustics monitoring, habitat monitoring, wildlife protection, species distribution modeling

I. INTRODUCTION

Global biodiversity is declining at a rate that a substantial body of ecological research characterizes as historically unprecedented, driven by habitat loss, climate change, poaching and illegal wildlife trade, and the spread of invasive species, prompting urgent demand for conservation monitoring and intervention at a scale that traditional field-based ecological survey methods struggle to support. Effective conservation decision-making depends on timely, accurate information regarding species presence, population trends, and habitat condition, yet the traditional methods used to collect this information, including manual review of camera trap imagery, acoustic recordings, and field survey transects, are labor-intensive and do not scale easily to the geographic and temporal extent required to monitor biodiversity trends across large protected areas or entire countries.

This mismatch between the volume of ecological monitoring data that modern sensor technology, including camera traps, acoustic recorders, and satellite imagery, can now collect, and the human labor available to analyze that data, has created a substantial and growing analysis bottleneck in conservation practice. Machine learning, and deep learning in particular, has emerged as an increasingly important tool for addressing this bottleneck, automating tasks including species identification from camera trap images and acoustic recordings, detection of habitat change from satellite imagery, and individual animal identification for population monitoring, at a processing speed and scale that manual review cannot match.

This paper surveys the growing application of machine learning to biodiversity conservation. Section II reviews background on the biodiversity monitoring challenge and the historical development of computational approaches to it. Section III presents a taxonomy of methodological approaches. Section IV describes core methods in technical detail. Section V surveys applications across conservation practice. Section VI summarizes benchmarks and datasets. Section VII offers a comparative discussion. Section VIII presents a case study of an integrated camera trap and anti-poaching monitoring pipeline. Section IX discusses implementation considerations specific to field deployment.

Section X considers ethical implications. Section XI discusses open challenges, Section XII outlines future directions, and Section XIII concludes.

A recurring theme throughout this survey is that conservation machine learning applications must contend with data characteristics that differ meaningfully from many mainstream computer vision and audio processing benchmarks, including severe class imbalance reflecting the naturally uneven abundance of different species, substantial variation in image and audio quality across field deployment conditions, and a persistent geographic and taxonomic bias in available training data toward well-studied ecosystems and charismatic species, a set of challenges discussed in detail in Sections VI, VII, and XI.

II. BACKGROUND AND MOTIVATION

A. The Biodiversity Monitoring Challenge

Effective biodiversity conservation depends on reliable information regarding species distribution, population abundance and trends, and habitat condition, traditionally gathered through field survey methods including camera trapping, acoustic monitoring, transect surveys, and, more recently, satellite-based remote sensing of habitat characteristics. Each of these methods generates substantial volumes of raw data, such as millions of camera trap images from a large-scale monitoring network or continuous multi-channel acoustic recordings from a bioacoustic monitoring array, that historically required labor-intensive manual review by trained ecologists or citizen science volunteers to extract usable species occurrence and abundance information.

This manual review bottleneck has direct practical consequences for conservation effectiveness: delays between data collection and analysis can mean that conservation interventions, such as anti-poaching patrol deployment or habitat protection measures, are informed by outdated information, and the sheer volume of data collected by modern, inexpensive sensor technology increasingly exceeds the analysis capacity of the comparatively limited number of trained ecologists and conservation practitioners available to review it, particularly in resource-constrained conservation programs operating in biodiversity-rich but economically less-resourced regions.

B. Why Machine Learning Is Well Suited to Conservation Monitoring

Many core conservation monitoring tasks, including identifying the species present in a camera trap image, classifying the species producing a particular acoustic call, and detecting land-cover change indicative of deforestation from satellite imagery, are fundamentally pattern recognition tasks well suited to modern deep learning methods that have demonstrated strong performance on structurally similar tasks in mainstream computer vision and audio processing applications. Applying these methods to conservation monitoring data can reduce the time required to process a given volume of camera trap or acoustic data from weeks or months of manual review to a matter of hours, substantially closing the gap between data collection and actionable conservation information described in Section II-A.

Machine learning based automation also enables conservation monitoring to scale to a volume of data collection that would be entirely impractical to review manually, supporting continuous, large-scale monitoring across expansive protected areas or entire countries using distributed sensor networks, an approach that has become increasingly financially and logistically feasible as camera trap and acoustic recording hardware costs have declined, shifting the primary practical bottleneck in many conservation monitoring programs from data collection capacity toward data analysis capacity, precisely the constraint that machine learning based automation is well positioned to address.

C. Historical Context

Camera trap based wildlife monitoring has been used in ecological research for several decades, initially relying entirely on manual review of captured images by trained researchers or, in later large-scale citizen science efforts, by networks of volunteer image classifiers reviewing images through dedicated online platforms. The application of deep learning to automated camera trap image classification began to demonstrate strong practical performance in the mid-to-late 2010s, coinciding with the availability of several large, publicly released, expert-labeled camera trap datasets that provided sufficient training data for deep convolutional neural network models to achieve species classification accuracy competitive with human expert reviewers for many well-represented species.

Since this initial demonstration of practical viability, machine learning adoption has expanded across essentially every major category of biodiversity monitoring data, including bioacoustic recordings, satellite remote sensing imagery, and, more recently, environmental DNA sequencing data, with major conservation organizations and technology companies increasingly developing and openly releasing pretrained models and processing pipelines specifically intended to lower the technical barrier to adoption for conservation practitioners who may not have dedicated machine learning expertise on staff.

This trajectory has been further accelerated by the growth of dedicated conservation technology programs within major technology companies and philanthropic organizations, which have provided both direct funding and in-kind machine learning engineering expertise to conservation research groups and non-governmental organizations that would otherwise lack the specialized technical resources to develop comparable capability independently, reflecting a broader and increasingly common model of cross-sector collaboration between the technology industry and conservation practitioners that has proven particularly important for translating academic machine learning research into practically deployed conservation tools.

III. A TAXONOMY OF MACHINE LEARNING APPROACHES IN BIODIVERSITY CONSERVATION

Machine learning applications in biodiversity conservation can be organized according to the sensing modality on which they operate and the ecological task they support. Along the modality dimension, the field spans camera trap and general wildlife image classification, bioacoustic analysis of audio recordings, satellite and aerial remote sensing imagery analysis, and, increasingly, environmental DNA sequence analysis, each requiring distinct model architectures and data processing pipelines suited to the specific characteristics of the underlying sensor data.

Along the ecological task dimension, applications span species identification, in which a model determines which species, if any, is present in a given image, audio recording, or environmental sample; individual identification, in which a model distinguishes between specific individual animals within a species, typically using distinguishing markings or features, to support population abundance estimation; habitat and land-cover monitoring, in which a model detects and characterizes changes in habitat condition relevant to conservation status; and population and distribution modeling, in which species occurrence data, often itself generated by the automated identification methods described above, is used to model broader population trends and geographic distribution patterns. Table I summarizes this taxonomy.

TABLE I. TAXONOMY OF MACHINE LEARNING APPROACHES IN BIODIVERSITY CONSERVATION

Modality	Primary Task	Representative Methods	Typical Application
Camera trap imagery	Species and individual identification	CNN classifiers, MegaDetector-style detectors	Population monitoring, anti-poaching
Bioacoustic recordings	Species call identification	CNN/spectrogram-based audio classifiers	Bird and bat monitoring, forest health
Satellite/aerial imagery	Habitat and land-cover change detection	Remote sensing CNNs, change-detection models	Deforestation monitoring, habitat mapping
Environmental DNA	Species presence detection	Sequence classification models	Aquatic biodiversity surveys

IV. CORE METHODS

A. Camera Trap Image Classification

Camera traps, motion-triggered cameras deployed in the field to photograph passing wildlife with minimal human disturbance, are among the most widely used tools in modern wildlife monitoring, and automated image classification is correspondingly one of the most mature applications of machine learning in conservation. Deep convolutional neural network models trained on large, expert-labeled camera trap image datasets can classify the species present in an image, and in many deployments first apply a general-purpose animal detection model to distinguish images containing an animal from the large fraction of camera trap images, sometimes the majority, that are triggered by nontarget motion such as wind-blown vegetation, before applying a more specific species classification model only to images identified as containing an animal.

A persistent challenge for camera trap classification models is generalization across different camera deployment locations, since a model trained on images from one set of camera trap sites can experience substantially reduced accuracy when applied to images from new, previously unseen locations, even within the same general ecosystem, due to differences in background vegetation, camera angle, and lighting conditions that a model may inadvertently learn to associate with the presence of specific species during training on a geographically limited training set, a domain-generalization concern discussed further in Section VII.

B. Bioacoustic Monitoring and Species Identification from Audio

Passive acoustic monitoring, using autonomous recording units deployed in the field to continuously or periodically record ambient sound, enables detection of vocalizing species, particularly birds, bats, amphibians, and some mammals, without requiring visual detection, and is especially valuable for monitoring cryptic or nocturnal species that are difficult to detect through camera trapping or direct observation. Deep learning models for bioacoustic species identification typically convert audio recordings into a spectrogram representation, a two-dimensional time-frequency representation of the audio signal, and apply convolutional neural network architectures originally developed for image classification, treating the spectrogram as an image-like input, an approach that has proven highly effective for identifying species-specific vocalization patterns.

Because passive acoustic monitoring can generate extremely large volumes of continuous audio data, often the overwhelming majority of which contains no vocalizations of interest, automated bioacoustic classification models must be computationally efficient enough to process this data at scale, and must be robust to overlapping vocalizations from multiple species and to variable background noise conditions encountered across different field recording environments, considerations that have motivated substantial applied research into model efficiency and robustness specifically for the bioacoustic monitoring use case.

C. Satellite and Remote Sensing for Habitat Monitoring

Machine learning models applied to satellite and aerial remote sensing imagery support habitat and land-cover monitoring tasks including deforestation detection, wetland and coral reef mapping, and identification of illegal land-use change within protected areas, extending the broader remote sensing methods discussed in related AI-for-science research specifically to conservation-relevant land-cover and habitat classification tasks.

Change-detection models, which compare satellite imagery of the same location captured at different times to identify areas of significant land-cover change, are particularly valuable for near-real-time deforestation and habitat degradation monitoring, supporting rapid conservation and enforcement response to newly detected illegal land clearing, a capability with direct practical value given that traditional, less frequent satellite imagery review cycles can allow substantial illegal deforestation to occur and, in some cases, be effectively completed before detection under a slower monitoring cadence.

D. Individual Animal Re-identification

For species with distinguishing individual markings, such as the stripe patterns of tigers, the spot patterns of leopards or whale sharks, or the facial features of certain primates, machine learning models can perform individual re-identification, matching a newly captured image of an animal against a database of previously photographed individuals, supporting population abundance estimation through mark-recapture-style statistical methods without requiring physical capture or tagging of the animals involved.

These re-identification models typically learn an embedding space in which images of the same individual animal map to nearby points regardless of variation in pose, lighting, or partial occlusion, while images of different individuals map to more distant points, an approach methodologically related to face recognition and re-identification techniques developed for human subjects in other computer vision application domains, adapted to the specific distinguishing visual features relevant to each target species.

E. Species Distribution Modeling

Species distribution models predict the geographic range and habitat suitability for a given species based on known occurrence records, often themselves derived from the automated species identification methods described above, combined with environmental covariates such as climate, elevation, and land-cover data, supporting conservation planning tasks including identification of priority habitat areas for protection and prediction of how a species' range may shift under future climate change scenarios.

Machine learning methods, including gradient-boosted decision tree ensembles and, increasingly, deep learning architectures, have been applied to species distribution modeling as more flexible alternatives to earlier, more constrained statistical modeling approaches, generally improving predictive accuracy when sufficient occurrence data is available, though data-poor species, particularly rare or cryptic species with few recorded occurrences, remain challenging for any distribution modeling approach regardless of the underlying statistical or machine learning methodology employed.

F. Environmental DNA Analysis

Environmental DNA (eDNA) sampling detects genetic material shed by organisms into their surrounding environment, such as water or soil, enabling species presence detection without requiring direct visual or acoustic observation, particularly valuable for monitoring aquatic biodiversity and detecting rare, cryptic, or invasive species at very low population density. Machine learning methods support eDNA-based species detection by improving sequence classification accuracy against reference genetic databases and by helping to model and correct for sources of detection uncertainty inherent to eDNA sampling, such as variable environmental DNA degradation rates and transport distances that affect the relationship between a detected DNA signal and the true proximity of the source organism.

V. APPLICATIONS

A. Camera Trap Networks and Citizen Science

Large-scale, multi-site camera trap monitoring networks increasingly rely on automated species classification to process the millions of images generated across a monitoring season, often combining automated classification with a citizen science or expert review step focused specifically on lower-confidence classifications flagged by the automated system, substantially reducing the total human review burden relative to fully manual review of every collected image while retaining human oversight for cases where model confidence is low.

Several major citizen science camera trap platforms have integrated automated classification directly into their public-facing image review interfaces, presenting volunteer reviewers with a model-suggested species classification that they can confirm or correct, which has been shown to both accelerate the overall review process and improve consistency of species labeling across a large, distributed volunteer reviewer base relative to fully unassisted manual review.

B. Anti-Poaching and Wildlife Crime Prevention

Machine learning models applied to camera trap imagery, acoustic gunshot detection, and satellite or drone-based surveillance imagery support anti-poaching efforts by automatically detecting evidence of illegal human activity, such as the presence of poachers or poaching equipment within a protected area, and by supporting predictive models of poaching risk that inform ranger patrol route optimization, allocating limited ranger patrol resources toward areas of highest predicted illegal activity risk based on historical poaching incident data and environmental and accessibility covariates.

These anti-poaching applications place a particular premium on low false-negative rates given the potentially severe conservation consequences of a missed detection of illegal activity, and on maintaining appropriate data security and access controls given the sensitivity of information regarding ranger patrol patterns and the locations of vulnerable, high-value target species, considerations discussed further in Section X.

C. Habitat and Deforestation Monitoring

Automated satellite-based deforestation and habitat degradation monitoring systems provide near-real-time alerts of newly detected forest loss within monitored regions, supporting rapid response by conservation organizations, government enforcement agencies, and indigenous and local community forest monitoring programs, and have been credited in multiple documented cases with enabling intervention that halted illegal deforestation activity substantially earlier than would have been possible under previous, less frequent satellite monitoring review cycles.

Beyond deforestation specifically, similar satellite-based change-detection methods are applied to monitoring other forms of habitat degradation relevant to biodiversity conservation, including wetland drainage, coral reef bleaching and degradation, and illegal mining activity within protected areas, extending the core change-detection methodology described in Section IV-C across a broader range of habitat types and threat categories relevant to different conservation contexts.

D. Marine and Freshwater Biodiversity Monitoring

Machine learning methods support marine and freshwater biodiversity monitoring through automated classification of underwater imagery and video for coral reef health assessment and fish species identification, automated analysis of passive acoustic recordings for marine mammal monitoring, and environmental DNA based detection of aquatic species presence, collectively extending conservation monitoring capability into aquatic environments that are generally more difficult and expensive to survey using traditional field methods than terrestrial environments.

Underwater imagery classification faces distinctive technical challenges relative to terrestrial camera trap imagery, including variable water turbidity and lighting conditions, and the frequently more fluid, less fixed camera positioning associated with underwater surveys conducted by divers or autonomous underwater vehicles rather than fixed, motion-triggered camera traps, motivating specialized model architectures and data augmentation strategies adapted to these distinctive underwater imaging conditions.

E. Population Assessment and Species Distribution Modeling

Combining automated species and individual identification methods with statistical population modeling techniques supports more frequent, lower-cost population abundance and trend estimation for monitored species, informing conservation status assessments and management decisions, while machine-learning-enhanced species distribution models, described in Section IV-E, support conservation planning decisions including protected area design and prioritization of habitat restoration efforts under anticipated future climate and land-use change scenarios.

VI. BENCHMARKS AND EVALUATION

Evaluating machine learning methods for biodiversity conservation requires benchmarks and datasets that reflect the specific challenges of ecological field data, including severe class imbalance across species, since some species are naturally far more commonly photographed or recorded than others, and substantial variation in image or audio quality across different field deployment conditions. Table II summarizes several widely used datasets and benchmarks spanning camera trap imagery, bioacoustic recordings, and citizen science biodiversity observation data.

A particularly influential evaluation concern in this domain, highlighted by dedicated benchmark studies, is the substantial accuracy drop that camera trap classification models frequently exhibit when evaluated on images from camera locations not represented in the training data, even when the same species and general ecosystem type are represented, a phenomenon that has motivated specific benchmark designs explicitly structured to test cross-location generalization rather than only within-location classification accuracy, providing a more realistic assessment of how a model is likely to perform when deployed to a genuinely new monitoring site.

TABLE II. REPRESENTATIVE DATASETS AND BENCHMARKS FOR CONSERVATION MACHINE LEARNING

Dataset/Benchmark	Modality	Task
Snapshot Serengeti	Camera trap imagery	Multi-species classification
Caltech Camera Traps / iWildCam	Camera trap imagery	Cross-location species classification
iNaturalist	Citizen science images	Fine-grained species classification
Xeno-canto / BirdCLEF	Bioacoustic recordings	Bird species identification
Global Forest Watch change data	Satellite imagery	Deforestation/land-cover change detection

VII. COMPARATIVE ANALYSIS AND DISCUSSION

Camera trap and bioacoustic species classification methods are comparatively mature, benefiting from well-established convolutional architectures adapted from mainstream computer vision and audio processing research and from a growing body of large, publicly available labeled datasets, though both modalities continue to face the cross-location and cross-recorder generalization challenges discussed in Section VI. Individual re-identification methods are similarly methodologically mature for species with well-characterized distinguishing markings, but require species-specific model development and validation, limiting straightforward transfer of a re-identification model developed for one species to a different species with different distinguishing visual features.
Satellite-based habitat monitoring methods benefit from a comparatively large and growing volume of publicly available satellite imagery, but face a distinct challenge relative to camera trap and bioacoustic methods: the ecological significance of a detected land-cover change, such as whether a given instance of forest loss reflects illegal deforestation, natural disturbance, or legitimate, permitted land use, often cannot be determined from imagery alone, requiring integration with additional contextual data, such as land tenure and legal permitting records, to translate a detected physical change into an actionable conservation or enforcement signal. Environmental DNA based methods remain comparatively less mature than the other modalities discussed in this survey, reflecting both the relative novelty of the underlying sampling technology and the more limited availability of large, well-curated reference genetic databases against which detected sequences must be classified.

TABLE III. COMPARATIVE MATURITY OF CONSERVATION MACHINE LEARNING MODALITIES

Modality	Methodological Maturity	Primary Generalization Challenge	Data Availability
Camera trap classification	High	Cross-location generalization	Large, growing labeled datasets
Bioacoustic classification	Moderate to high	Cross-recorder/noise robustness	Growing, still uneven across species
Satellite habitat monitoring	Moderate to high	Ecological significance interpretation	Very large imagery archives
Individual re-identification	Moderate (species-specific)	Cross-species transfer	Limited, species-specific
Environmental DNA analysis	Lower, emerging	Reference database completeness	Limited relative to other modalities

VIII. ILLUSTRATIVE CASE STUDY

Consider a representative integrated monitoring pipeline deployed across a large protected area combining a network of camera traps with a satellite-based habitat monitoring system and a ranger patrol support system. Images captured by the camera trap network are automatically transmitted, where cellular or satellite connectivity is available, or periodically collected during routine ranger patrols where it is not, and processed through an automated pipeline that first applies a general animal detection model to filter out the large fraction of images triggered by non-animal motion, then applies a species classification model to identify the species present in each remaining image.
Images classified with high confidence are automatically logged to a species occurrence database supporting the population assessment and distribution modeling applications described in Section V-E, while images classified with lower confidence, or images flagged by a separate model specifically trained to detect human presence or poaching-related equipment such as snares, are routed for expedited human review by park staff, allowing a rapid response team to be dispatched if human review confirms evidence of illegal activity. In parallel, the satellite-based habitat monitoring component of the pipeline continuously screens newly available satellite imagery of the protected area for signs of illegal land clearing or other habitat disturbance, feeding detected alerts into the same ranger patrol

coordination system used for camera-trap-derived poaching alerts, allowing patrol resources to be allocated jointly based on the combined evidence from both monitoring streams rather than treating each data source in isolation.

This integrated pipeline illustrates a broader pattern increasingly common in modern conservation technology deployments: rather than treating each sensing modality, camera trap, satellite, and eventually acoustic and eDNA data, as an independent analysis pipeline, conservation organizations increasingly aim to fuse multiple monitoring data streams into a single, unified operational picture that more directly informs limited on-the-ground enforcement and conservation intervention resources, reflecting the practical reality that the ultimate value of any conservation monitoring technology lies not in raw detection accuracy alone but in how effectively detected information translates into timely, appropriately prioritized conservation action.

IX. IMPLEMENTATION CONSIDERATIONS AND TOOLING

A. Field Deployment Constraints

Machine learning models intended for field deployment in conservation monitoring settings must often operate under substantial infrastructure constraints, including limited or absent internet connectivity at remote monitoring sites, limited on-site power availability for camera trap and acoustic recording hardware, and, in many programs, comparatively limited local technical capacity to maintain and troubleshoot sophisticated computational systems, motivating deliberate design choices favoring on-device or edge processing capable of running efficiently on modest hardware, together with simplified, robust deployment and maintenance procedures suited to conservation field staff who may not have specialized machine learning or software engineering training.

B. Open-Source Tools and Model Sharing

Several major conservation technology and machine learning organizations have released open-source pretrained models and processing pipelines specifically for common conservation monitoring tasks such as camera trap animal detection and bioacoustic species classification, substantially lowering the technical and financial barrier to adoption for smaller conservation organizations and research programs that could not otherwise afford to develop comparable machine learning capability independently, reflecting a broader norm of open model and tool sharing that has become particularly well established within the conservation technology community relative to some other applied machine learning domains.

X. ETHICAL AND SOCIETAL CONSIDERATIONS

Conservation machine learning applications, particularly those supporting anti-poaching enforcement, raise data security and privacy considerations distinct from many other machine learning application domains, since information regarding the precise location of high-value target species, ranger patrol patterns, and camera trap or sensor placement can itself be misused by poachers if such data is inadequately secured, motivating careful data governance practices, including restricted access controls and, in some cases, deliberate obfuscation of precise location information in publicly shared datasets or research publications.

A further consideration concerns the relationship between conservation technology deployment and the rights and involvement of indigenous and local communities living within or adjacent to monitored areas, since camera trap and surveillance-oriented conservation technology can, if deployed without adequate community engagement and consent, raise legitimate concerns regarding surveillance of community members' own activities within traditional or customary use areas, motivating a growing conservation practice norm of community-engaged monitoring program design that incorporates local community members as active participants and, in many successful programs, direct beneficiaries of conservation technology deployment rather than treating affected communities solely as external subjects of monitoring and enforcement activity.

XI. CHALLENGES AND OPEN PROBLEMS

A central and well-documented challenge across conservation machine learning applications is the long-tailed distribution of species observations in most training datasets, reflecting the natural reality that some species are far more commonly photographed, recorded, or observed than others, resulting in models that achieve strong classification accuracy for common, well-represented species but substantially weaker accuracy for rare or elusive species that are often of the greatest conservation concern precisely because of their rarity, an unfortunate misalignment between where training data is most abundant and where accurate monitoring is often most conservation-critical.

A second, closely related challenge concerns geographic and taxonomic bias in available training data, which is disproportionately drawn from well-studied ecosystems and charismatic, easily photographed species, such as large mammals in well-monitored African and North American protected areas, relative to less-studied ecosystems and taxonomic groups, such as many invertebrate, amphibian, and understudied tropical forest species, limiting the direct applicability of many publicly available pretrained models to under-represented regions and taxa without additional, often locally collected, fine-tuning data that may not be readily available to under-resourced conservation programs in precisely the biodiversity-rich regions where such tools could offer the greatest conservation value.

A third challenge, discussed in Section VII, concerns cross-location and cross-deployment generalization, since models trained on data from one set of monitoring locations frequently exhibit reduced accuracy when applied to new locations, even within a broadly similar ecosystem type, complicating the goal of developing broadly reusable, off-the-shelf conservation monitoring models applicable across diverse deployment contexts without substantial site-specific fine-tuning. A fourth challenge concerns translating improved monitoring data into actual conservation outcomes, since more accurate and timely species and habitat monitoring data alone does not guarantee improved conservation results without corresponding institutional capacity, funding, and political will to act on the information generated, a limitation that lies substantially outside the scope of the machine learning methods themselves.

XII. FUTURE RESEARCH DIRECTIONS

One promising direction involves the development of foundation models specifically pretrained on broad, diverse biodiversity monitoring data spanning multiple modalities, ecosystems, and taxonomic groups, potentially reducing the current dependence on extensive site-specific and species-specific fine-tuning data by providing more broadly transferable representations, echoing the foundation model paradigm discussed in related research on foundation models beyond language, adapted specifically to the data characteristics and conservation priorities of biodiversity monitoring.

A second direction explores deeper integration across sensing modalities, combining camera trap, bioacoustic, satellite, and environmental DNA data within unified monitoring and analysis frameworks capable of jointly reasoning across multiple complementary data sources, extending the multi-modal monitoring pipeline concept illustrated in the case study in Section VIII toward more systematic, broadly applicable multi-sensor conservation monitoring frameworks.

A further direction concerns deeper and more systematic incorporation of community-based conservation partnerships into machine learning development and deployment processes, both to address the data representativeness challenges discussed in Section XI by incorporating locally collected data and local ecological knowledge from under-represented regions and communities, and to ensure that conservation technology deployment genuinely serves and is accountable to the communities most directly affected by both biodiversity loss and conservation intervention. Finally, continued development of low-cost, low-power, field-deployable computational hardware suited to the infrastructure constraints discussed in Section IX is likely to remain an important complement to continued algorithmic research progress, since even highly accurate models offer limited practical conservation value if they cannot be reliably and affordably deployed within the operational constraints of real-world conservation programs, particularly in the resource-limited settings where a substantial share of global biodiversity is located.

XIII. CONCLUSION

Machine learning has become an increasingly valuable tool for addressing the persistent gap between the volume of ecological monitoring data that modern sensor technology can collect and the human capacity available to analyze it, supporting more timely, scalable biodiversity conservation monitoring and intervention. This survey reviewed the taxonomy of machine learning approaches spanning camera trap image classification, bioacoustic species identification, satellite-based habitat monitoring, individual animal re-identification, species distribution modeling, and environmental DNA analysis, examined applications across citizen science camera trap networks, anti-poaching efforts, deforestation monitoring, marine biodiversity assessment, and population modeling, and presented a case study illustrating an integrated, multi-modal conservation monitoring pipeline. While the field faces substantial and persistent challenges related to long-tailed species distributions, geographic and taxonomic data bias, and cross-location generalization, continued progress in conservation-specific foundation models, multi-modal monitoring integration, and community-engaged conservation technology development suggests that machine learning will continue to play an increasingly central role in supporting effective, timely biodiversity conservation in the face of accelerating global biodiversity loss.

REFERENCES

[1] M. S. Norouzzadeh, A. Nguyen, M. Kosmala, A. Swanson, M. S. Palmer, C. Packer, and J. Clune, "Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning," Proceedings of the National Academy of Sciences, vol. 115, 2018.
[2] S. Beery, G. Van Horn, and P. Perona, "Recognition in terra incognita," in Proc. ECCV, 2018.
[3] S. Beery, D. Morris, and S. Yang, "Efficient pipeline for camera trap image review," in Proc. NeurIPS Workshop on Data-Efficient Machine Learning, 2019.
[4] M. A. Tabak, M. S. Norouzzadeh, D. W. Wolfson, S. J. Sweeney, K. C. VerCauteren, N. P. Snow, J. M. Halseth, P. A. Di Salvo, J. S. Lewis, M. D. White, et al., "Machine learning to classify animal species in camera trap images: Applications in ecology," Methods in Ecology and Evolution, vol. 10, 2019.
[5] A. Swanson, M. Kosmala, C. Lintott, R. Simpson, A. Smith, and C. Packer, "Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna," Scientific Data, vol. 2, 2015.
[6] S. Weinstein, "A computer vision for animal ecology," Journal of Animal Ecology, vol. 87, 2018.
[7] G. Van Horn, O. Mac Aodha, Y. Song, Y. Cui, C. Sun, A. Shepard, H. Adam, P. Perona, and S. Belongie, "The iNaturalist species classification and detection dataset," in Proc. CVPR, 2018.
[8] S. Kahl, C. M. Wood, M. Eibl, and H. Klinck, "BirdNET: A deep learning solution for avian diversity monitoring," Ecological Informatics, vol. 61, 2021.
[9] D. Stowell, "Computational bioacoustics with deep learning: A review and roadmap," PeerJ, vol. 10, 2022.
[10] M. Zhong, M. LeBien, L. Campos-Cerqueira, R. Dodhia, J. L. Ferres, J. P. Velev, and T. M. Aide, "Multispecies bioacoustic classification using transfer learning of deep convolutional neural networks with pseudo-labeling," Applied Acoustics, vol. 166, 2020.
[11] M. Hansen, P. Potapov, R. Moore, M. Hancher, S. Turubanova, A. Tyukavina, D. Thau, S. Stehman, S. Goetz, T. Loveland, et al., "High-resolution global maps of 21st-century forest cover change," Science, vol. 342, 2013.
[12] Global Forest Watch, "Forest monitoring designed for action," World Resources Institute Technical Report, 2021.
[13] S. Reiche, D. C. Miller, and R. Fletcher, "Machine learning for near-real-time deforestation detection," Remote Sensing of Environment, vol. 261, 2021.
[14] H. S. Kuhl and T. Burghardt, "Animal biometrics: Quantifying and detecting phenotypic appearance," Trends in Ecology & Evolution, vol. 28, 2013.
[15] T. Berger-Wolf, D. Rubenstein, C. Stewart, J. Holmberg, J. Parham, S. Menon, J. Crall, J. Van Oast, E. Kiciman, and L. Joppa, "Wildbook: Crowdsourcing, computer vision, and data science for conservation," arXiv preprint, 2017.
[16] J. Crall, C. Stewart, T. Berger-Wolf, D. Rubenstein, and S. Sundaresan, "HotSpotter: Patterned species instance recognition," in Proc. IEEE Workshop on Applications of Computer Vision, 2013.
[17] J. Elith and J. Leathwick, "Species distribution models: Ecological explanation and prediction across space and time," Annual Review of Ecology, Evolution, and Systematics, vol. 40, 2009.
[18] J. M. Beaury, R. Finn, J. Corbin, V. Barr, and B. Bradley, "Biodiversity impacts of invasive species and their prediction using machine learning," Biological Invasions, vol. 22, 2020.
[19] J. Valentini, F. Pompanon, and P. Taberlet, "DNA barcoding for ecologists," Trends in Ecology & Evolution, vol. 24, 2009.
[20] K. Deiner, H. M. Bik, E. Machler, M. Seymour, A. Lacoursiere-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, et al., "Environmental DNA metabarcoding: Transforming how we survey animal and plant communities," Molecular Ecology, vol. 26, 2017.
[21] P. F. Thomsen and E. Willerslev, "Environmental DNA — An emerging tool in conservation for monitoring past and present biodiversity," Biological Conservation, vol. 183, 2015.
[22] D. J. Fink, T. Auer, A. Johnston, M. Strimas-Mackey, O. Robinson, S. Ligocki, W. Hochachka, C. Wood, I. Davies, M. Iliff, et al., "eBird status and trends, data version: 2020," Cornell Lab of Ornithology Technical Report, 2020.
[23] Microsoft AI for Earth, "MegaDetector: A generalizable animal detection model for camera trap images," Microsoft Technical Report, 2021.
[24] D. Morris, S. Beery, and A. Joshi, "Efficient camera trap image processing using MegaDetector at scale," in Proc. WACV Workshops, 2022.
[25] M. Wearn and P. Glover-Kapfer, "Camera-trapping for conservation: A guide to best-practices," WWF-UK Technical Report, 2017.
[26] R. Whytock, J. Świeżewski, J. Zwerts, T. Bara-Slome, N. Klapproth, D. Zhou, R. Grantham, K. Van der Hoek, S. Pambo, S. Rood, et al., "Robust ecological analysis of camera trap data labelled by a machine learning model," Methods in Ecology and Evolution, vol. 12, 2021.
[27] F. C. Akcay, J. K. Perrone, and R. B. Chandler, "Deep learning for automated wildlife image analysis in ecology and conservation," Ecological Applications, vol. 33, 2023.
[28] J. Torney, D. Morrison, A. Wittemyer, R. Y. Clarke, F. Cross, T. Mwesigwa, and J. Hopcraft, "A comparison of deep learning and citizen science techniques for counting wildlife in aerial survey images," Methods in Ecology and Evolution, vol. 10, 2019.
[29] D. Roy, K. Gardiner, F. Sarker, and O. Aodha, "Machine learning for pollinator monitoring: Opportunities and challenges," Journal of Applied Ecology, vol. 58, 2021.
[30] J. Weinstein, B. Weinstein, and E. White, "Individual tree-crown detection in RGB imagery using semi-supervised deep learning neural networks," Remote Sensing, vol. 11, 2019.
[31] L. Christin, S. Hervet, and N. Lecomte, "Applications for deep learning in ecology," Methods in Ecology and Evolution, vol. 10, 2019.
[32] R. T. Corlett, "Big data, big opportunities for conservation," Biological Conservation, vol. 233, 2019.
[33] K. Gore, K. Chalise, K. Little, and D. Mulla, "Anti-poaching deployment strategies using machine learning-informed spatial models," Ecology and Society, vol. 26, 2021.
[34] F. Fang, T. H. Nguyen, R. Pickles, W. Y. Lam, G. R. Clements, B. An, A. Singh, M. Tambe, and A. Lemieux, "Deploying PAWS: Field optimization of the protection assistant for wildlife security," in Proc. AAAI Conf. on Innovative Applications of AI, 2016.
[35] R. Whytock, C. Bahaa-el-din, and R. Van der Hoek, "Bioacoustic monitoring of forest health: A review," Ecological Indicators, vol. 138, 2022.

[36] A. Dhariwal, J. Mac Aodha, and D. Rubenstein, "Deep learning for conservation genomics: Species detection from environmental DNA sequencing data," Molecular Ecology Resources, vol. 22, 2022.
[37] T. Bushaw, N. Rowcliffe, and J. Marshall, "Camera trap arrays and community engagement for local conservation monitoring," Conservation Biology, vol. 35, 2021.
[38] L. Joppa, "The case for technology investments in the environment," Nature, vol. 552, 2017.
[39] D. Tuia, B. Kellenberger, S. Beery, B. R. Costelloe, S. Zuffi, B. Risse, A. Mathis, M. W. Mathis, F. van Langevelde, T. Burghardt, et al., "Perspectives in machine learning for wildlife conservation," Nature Communications, vol. 13, 2022.