



OPEN ACCESS

EDITED AND REVIEWED BY
Des B. A. Thompson,
Edinburgh, United Kingdom

*CORRESPONDENCE

Yizhar Lavner

✉ yizharle@telhai.ac.il

Cristian Pérez-Granados

✉ cristian.perez@ua.es

RECEIVED 27 October 2024

ACCEPTED 06 December 2024

PUBLISHED 23 December 2024

CITATION

Lavner Y and Pérez-Granados C (2024)

Editorial: Computational bioacoustics and automated recognition of bird vocalizations: new tools, applications and methods for bird monitoring.

Front. Bird Sci. 3:1518077.

doi: 10.3389/fbirs.2024.1518077

COPYRIGHT

© 2024 Lavner and Pérez-Granados. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Computational bioacoustics and automated recognition of bird vocalizations: new tools, applications and methods for bird monitoring

Yizhar Lavner^{1*} and Cristian Pérez-Granados^{2,3*}

¹Department of Computer Science, Tel-Hai College, Upper Galilee, Israel, ²Ecology Department, Alicante University, Alicante, Spain, ³Conservation Biology Group, Landscape Dynamics and Biodiversity Program, Forest Science and Technology Center of Catalonia (CTFC), Lleida, Spain

KEYWORDS

machine learning, deep learning, passive acoustic monitoring, bioacoustics, long-term bird monitoring, biodiversity, bird vocalizations, audio signal processing

Editorial on the Research Topic

[Computational bioacoustics and automated recognition of bird vocalizations: new tools, applications and methods for bird monitoring](#)

In recent decades, technological advancements have significantly transformed wildlife monitoring methods, particularly through the adoption of automated and non-invasive techniques like passive acoustic monitoring (PAM). PAM involves deploying sound recorders in the field to capture audio during specified times, which are followed by sound interpretation to gather information on the presence and behavior of species. A major challenge associated with PAM is the vast amount of audio data generated, necessitating the development of automated processing techniques utilizing machine learning, deep learning, or other advanced audio signal processing methods (Stowell, 2022). Birds have been the primary focus of PAM studies, making our understanding of how to effectively monitor and automatically identify bird species from audio recordings well-established. Indeed, PAM has been shown to be a reliable method for estimating bird species richness and population densities from sound recordings (Darras et al., 2019; Pérez-Granados and Traba, 2021). Additionally, it has been employed to investigate various ecological questions, including the detection of previously unnoticed habitat changes and shifts in biodiversity patterns (Ross et al., 2023).

However, despite these advances, there remain many challenges and uncertainties that need to be addressed, such as improving automated species identification accuracy, model evaluation, or pairing PAM with other automated techniques. Through this Research Topic, we present recent advancements in tools, methods, and algorithms for automated bird identification to enhance bird monitoring programs utilizing PAM.

One of the most important tasks in computational bioacoustics is tracking and monitoring biodiversity changes, which have been greatly advanced using PAM. Among the factors influencing biodiversity, agriculture plays a major role. Employing PAM to

assess the impact of agricultural practices on species occupancy and diversity requires careful planning and consideration of key factors for successful research. The work of [Molina-Mora et al.](#) discusses these factors and use PAM for assessing the impact of agricultural management practices on biodiversity, using birds as indicators. Specifically, their research focuses on coffee-growing areas in Costa Rica, comparing the effects of pruning and pesticide use over two years. Utilizing a mobile application for community-driven bioacoustic annotations to identify the vocalizations of selected species, they report that pruning negatively affects some species, while pesticide application reduces vocal activity and presumably the presence of all species studied.

Currently, most of PAM projects use automated signal recognition software and machine learning models to detect their species of interest ([Kahl et al., 2021](#); [Cole et al., 2022](#); [Pérez-Granados, 2023](#); [Lavner et al., 2024](#); [Sethi et al., 2024](#)). Yet, the application of machine learning models faces significant challenges due to the unique characteristics of bioacoustic data and the complexity of real-world deployment. For instance, PAM recordings often have low SNR, as bird vocalizations are embedded in substantial background noise. There is also a mismatch between training data, typically made with directional microphones, and field-recorded PAM data, which captures multiple species simultaneously, often with background noise. Additional challenges include class imbalance, distribution shifts, and inconsistent annotation quality, all of which hinder machine learning models' ability to generalize across varied conditions.

Two approaches are presented in the Research Topic to address these challenges: specialized models trained on specific datasets and tailored to answer particular tasks or specific research questions, and generalizable models capable of adapting to a variety of datasets and tasks.

A specialized solution is offered by [Haley et al.](#) who track the activity of mountain chickadee (*Poecile gambeli*) in North America, comparing a custom CNN model ([Madhusudhana et al., 2021](#)) to BirdNET ([Khal et al., 2021](#)). Their custom model, trained exclusively on PAM recordings, exhibited better performance and greater resilience to noise than BirdNET model, which highlights that specific models trained on specialized (i.e. locally trained) datasets may offer an advantage by being more closely aligned with the audio recording data.

In construct, [van Merriënboer et al.](#) advocate for the construction of bioacoustic generalizable or foundation models ("large scale machine learning models that are trained on broad data that can be adapted to a wide variety of downstream tasks", [Bommasani et al., 2021](#)), designed to learn from a specific training set or even a small number of examples, [few shot learning ([Nolasco et al., 2023](#); [Ghani et al., 2023](#))] generalize to new species, environments, and tasks beyond those they were initially trained on. They provide a comprehensive review of the challenges associated with developing such models, as well as the methods and metrics used for their evaluation.

Another challenge in species classification for PAM is the common practice of setting a threshold to determine whether a

species' vocalization is present in an audio clip. Typically, classification scores above the threshold are counted as detections ([Pérez-Granados, 2023](#)). Such selection affects the number of false positives or negatives, which can vary across different datasets or species, leading to suboptimal classifications. To resolve this issue, [Navine et al.](#) developed a 'threshold-free' bioacoustics analysis framework, which directly estimates call density—the proportion of detection windows containing the target vocalizations—regardless of classifier score, which may be applied to binary detection classifiers with fixed-size windows. A validation method estimates call density in a dataset, and probability distributions are generated for confidence scores in both positive and negative classes. These distributions are used to predict call densities at the site level, accounting for potential shifts in data distribution. Testing on real-world recordings of Hawaiian birds demonstrates this approach's robustness to variations in call density and classifier performance.

In PAM surveys it is difficult to estimate the precise location or the identity of vocalizing individuals, though such information is useful for studying birds' behavior or community dynamics. [Guggenberger et al.](#) propose a solution to this problem by coupling a video camera to an acoustic recorder equipped with a 64 microphone array, that enables to localize dozens of individuals, with high temporal and spatial resolution. To demonstrate its potential, they recorded and localized Arabian babblers (*Argya squamiceps*) during snake-mobbing events. The tagged birds allowed for precise identification of callers and their vocal timing, enabling the reconstruction of vocal social networks. The analysis revealed a periodic pattern in vocalizations, with age-specific inter-call duration, suggesting that snake-mobbing behavior is a learned activity within the group.

This topic highlights the potential of PAM for bird monitoring, including studying shifts in bird communities due to human land use, analyzing birds' social networks, and improving the development and assessment of robust, generalizable machine learning models. A significant challenge in advancing automated bird detection models is the scarcity of annotated PAM datasets that span diverse habitats and geographic regions worldwide. To address this, the bioacoustic research community should collaborate to create a global, standardized, and annotated PAM dataset. By pooling resources and adhering to common protocols, researchers can promote the development of new, comparable methods for passive bird surveys through the construction and use of scalable foundation models for bioacoustic monitoring, similar to successful applications in other fields ([O'Neill et al., 2023](#)). Such efforts would not only advance species identification and acoustic monitoring but also contribute significantly to global nature conservation.

Author contributions

YL: Conceptualization, Writing – original draft, Writing – review & editing. CP: Conceptualization, Writing – original draft, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

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. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., et al. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*. doi: 10.48550/arXiv.2108.07258
- Cole, J. S., Michel, N. L., Emerson, S. A., and Siegel, R. B. (2022). Automated bird sound classifications of long-duration recordings produce occupancy model outputs similar to manually annotated data. *Ornithol. Appl.* 124, duac003. doi: 10.1093/ornithapp/duac003
- Darras, K., Batáry, P., Furnas, B. J., Grass, I., Mulyani, Y. A., and Tschardtke, T. (2019). Autonomous sound recording outperforms human observation for sampling birds: a systematic map and user guide. *Ecol. Appl.* 29, e01954. doi: 10.1002/eap.v29.6
- Ghani, B., Denton, T., Kahl, S., and Klink, H. (2023). Global birdsong embeddings enable superior transfer learning for bioacoustic classification. *Sci. Rep.* 13, 22876. doi: 10.1038/s41598-023-49989-z
- Kahl, S., Wood, C. M., Eibl, M., and Klink, H. (2021). BirdNET: A deep learning solution for avian diversity monitoring. *Ecol. Inform.* 61, 101236. doi: 10.1016/j.ecoinf.2021.101236
- Lavner, Y., Melamed, R., Bashan, M., and Vortman, Y. (2024). The bioacoustic soundscape of a pandemic: continuous annual monitoring using a deep learning system in Agmon Hula Lake Park. *Ecol. Inf.* 80, 2024. doi: 10.1016/j.ecoinf.2024.102528
- Madhusudhana, S., Shiu, Y., Klinck, H., Fleishman, E., Liu, X., Nosal, E. M., et al. (2021). Improve automatic detection of animal call sequences with temporal context. *J. R. Soc. Interface* 18, 20210297. doi: 10.1098/rsif.2021.0297
- Nolasco, I., Singh, S., Morfi, V., Lostanlen, V., Strandburg-Peshkin, A., Vidaña-Vila, E., et al. (2023). Learning to detect an animal sound from five examples. *Ecol. Inf.* 77, 102258. doi: 10.1016/j.ecoinf.2023.102258
- O'Neill, A., Rehman, A., Gupta, A., Maddukuri, A., Gupta, A., Padalkar, A., et al. (2023). Open x-embodiment: Robotic learning datasets and rt-x models. *arXiv preprint arXiv:2310.08864*. doi: 10.48550/arXiv.2310.08864
- Pérez-Granados, C. (2023). BirdNET: applications, performance, pitfalls and future opportunities. *Ibis* 165, 1068–1075. doi: 10.1111/ibi.v165.3
- Pérez-Granados, C., and Traba, J. (2021). Estimating bird density using passive acoustic monitoring: a review of methods and suggestions for further research. *Ibis* 163, 765–783. doi: 10.1111/ibi.v163.3
- Ross, S. R. J., O'Connell, D. P., Deichmann, J. L., Desjonquères, C., Gasc, A., Phillips, J. N., et al. (2023). Passive acoustic monitoring provides a fresh perspective on fundamental ecological questions. *Funct. Ecol.* 37, 959–975. doi: 10.1111/1365-2435.14275
- Sethi, S. S., Bick, A., Chen, M. Y., Crouzeilles, R., Hillier, B. V., Lawson, J., et al. (2024). Large-scale avian vocalization detection delivers reliable global biodiversity insights. *Proc. Natl. Acad. Sci.* 121, e2315933121. doi: 10.1073/pnas.2315933121
- Stowell, D. (2022). Computational bioacoustics with deep learning: a review and roadmap. *PeerJ* 10, e13152. doi: 10.7717/peerj.13152