Bioacoustics on Tiny Hardware at the BioDCASE 2025 Challenge

Giovanni Carmantini,[1a] *Yasmine Benhamadi,*[2b] *Matthieu Carreau,*[2c] *Minkyung Kwak,*[3b] *Ilaria Morandi,*[4b]
Friedrich Förstner,[1,5a] *Pierre-Emmanuel Hladik,*[2c] *Mathieu Lagrange,*[2bc] *Pavel Linhart,*[6b] *Tereza Petrusková,*[3b]
Vincent Lostanlen,[2bc] *Stefan Kahl*[5,7]

[1] fold ecosystemics, 2 Bis Route d'Hadigny, 88330, Badménil-aux-Bois, France
[2] Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France
[3] Department of Ecology, Faculty of Science, Charles University, Viničná 7, Prague 2 12844, Czechia
[4] Department of Zoology, Faculty of Science, University of South Bohemia,
Branišovská 1716/31c, České Budějovice 37005, Czechia
[5] Chemnitz University of Technology, Straße der Nationen 62, 09111 Chemnitz, Germany
[6] Center for Sustainable Landscapes under Global Change (SustainScapes), Department of Biology,
Aarhus University, Ny Munkegade 114-116, DK-8000 Aarhus C, Denmark
[7] K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University,
159 Sapsucker Woods Road, Ithaca, NY 14850, USA

***Abstract*—The BioDCASE initiative aims to encourage the invention of methods for detection and classification of acoustic scenes and events (DCASE) within the domain of bioacoustics. We have contributed to the first edition of the BioDCASE challenge by means of a task named "bioacoustics on tiny hardware". The motivation for this task resides in the growing need for operating bioacoustic event detection algorithms on low-power autonomous recording units (ARU's). Participants were tasked with developing a detector of bird vocalizations from the yellowhammer *(Emberiza citrinella)*, given two hours of audio as a training set. The detector had to run within the resource constraints of an ESP32-S3 microcontroller unit. By evaluating the submitted models on a withheld dataset, we conducted an independent benchmark that assessed both classification performance and resource efficiency through multiple metrics: average precision, inference time, and memory usage. Our reported results confirm that recent advances in "tiny machine learning" (TinyML) have transformative potential for computational bioacoustics. For more information, please visit: https://biodcase.github.io/challenge2025/task3**

***Index Terms*—Acoustic event detection, autonomous recording units, bioacoustics, edge computing, passive acoustic monitoring.**

1. INTRODUCTION

Bioacoustics, understood as the science of sonic interaction in and between animals, requires sensors for data collection [1]. For animal behavior research [2] or biodiversity surveys [3], these sensors are typically deployed onto remote locations, off the electrical grid. As of today, most commercially available sensors for birds and land mammals are battery-powered, with a battery life of 200–500 hours and a cost of $100–$700 [4]. They record digital audio, either continuously or according to an intermittent schedule, and store it onto an SD card. Although they are often referred to as "autonomous recording units" (ARU), they are not fully autonomous: indeed, frequent round trip from the lab to the deployment site are necessary, in order to replace batteries, transfer the recorded data, and reset the SD card.

Previous publications have alerted on the lack of autonomy of current-generation ARU's and its negative consequences [5]–[7]:

1) An insufficient uptime or ill-suited schedule may cause the ARU to misportray these dynamics of vocal activity, eventually reducing its usefulness for statistical hypothesis testing [8].
2) The cost of battery replacement and the weight of hardware is a serious challenge for practitioners [9], [10].
3) Although they do not require direct contact or manipulation of the animals, ARU's may still be considered invasive if its maintenance raises the level of anthropogenic pressure; that is, the stress induced by the presence of humans [11].
4) The same data which are collected for research may be used for surveillance; an effect known as surveillance creep [12]. Some terrestrial bioacoustic datasets contain voices from people who are unaware of being recorded [13]. In the age of automated audio content analysis, the surveillance creep of acoustic sensors is not only a risk but a documented reality [14].
5) As technological advancements accelerate and costs decrease, the ecological impact of production and the growing issue of electronic waste (e-waste) have emerged as critical sustainability challenges [15]. For lack of a better end-of-life management [16], it would be prudent to strive for a lower dependency on batteries, or even switch to batteryless hardware [17].

Facing the drawbacks of current-generation ARU's, we propose to explore an alternative design: on-device sound event detection (SED). The key idea is to develop an algorithm which is able to analyze the audio stream in real-time and record sound event selectively. In the context of wireless acoustic sensor networks, this is known as edge computing as opposed to cloud computing. The main argument in favor of edge computing on ARU's is that the bioacoustic events of interest are typically few, while storing audio permanently is costly in terms of energy. If SED can be made efficient enough, the energy savings of selective storage can cover and outweigh the energy expenditure of edge computing [18]. Beyond the net gain in uptime, this simple idea has the potential to make the next generation of ARU's cheaper to maintain, less invasive, more privacy-preserving, and more durable.

Machine learning algorithms allow to automate various bioacoustics tasks, including call detection or species classification. Previous works demonstrate the feasibility of running these algorithms on "tiny hardware" such as microcontroller units (MCU's). For example, [19] have trained a MFCC-based classifier for species-specific call detection on a low-power MCU, with a memory usage of the order of

[a]The development of the evaluation system was funded by the German Federal Ministry of Education and Research as part of the "BirdNET+" project (FKZ 01-S22072). The competition received additional support from the German Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection through the "DeepBirdDetect" project (FKZ 67KI31040E).
[b]Supported by EU MSCA Doctoral Network Bioacoustic AI (101071532).
[c]This research was funded, in whole or in part, by the French National Research Agency (ANR) under the project OWL "ANR-23-IAS3-0003-01".

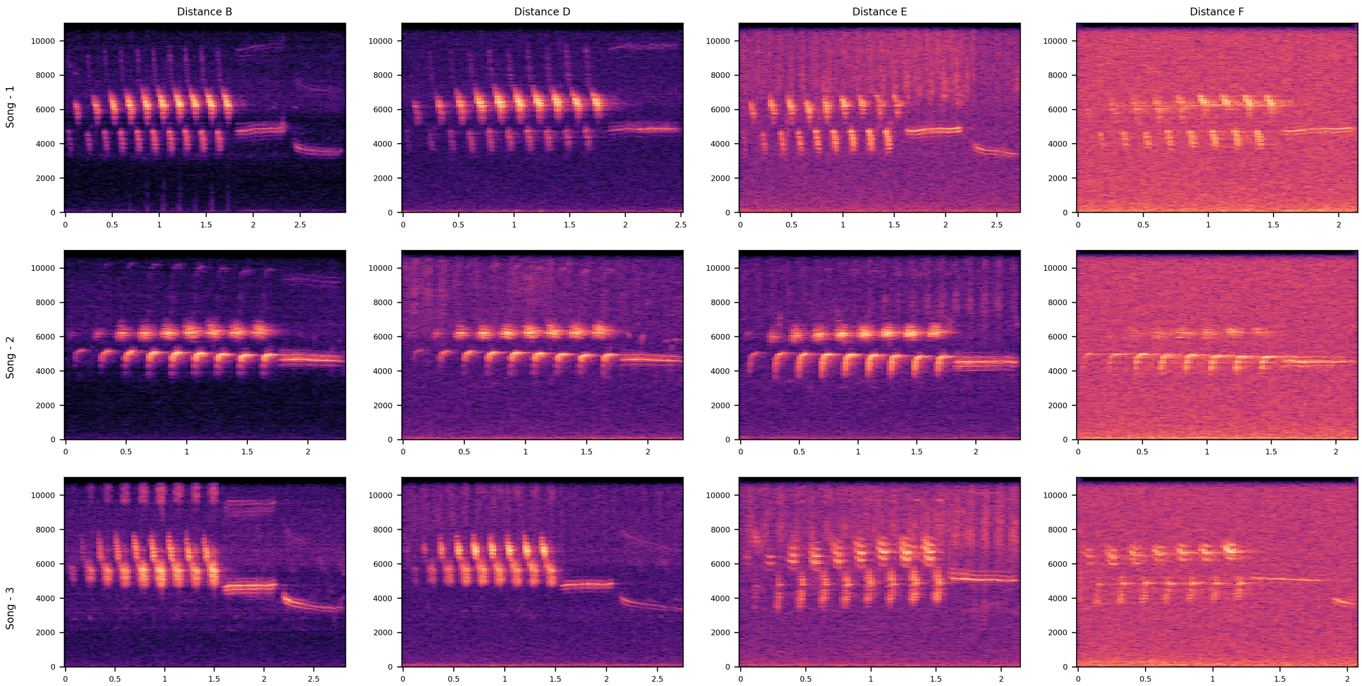

Fig. 1: Spectrogram representations of yellowhammer (*Emberiza citrinella*) song. Brighter colors denote relatively larger time–frequency magnitudes after normalization of visual contrast. Rows illustrate inter-individual variation while columns illustrate the effect of distance. See Section 2 for details.

one kilobyte. Another example is [20], who have shown the feasibility of SED on the AudioMoth, a widespread and open-source ARU [21]: the proposed algorithms rely on MFCC and Goertzel filtering and are applied to cicadas and gunshots.

Recent work at the intersection of "tiny machine learning" (TinyML) and bioacoustics has proposed to embed deep neural networks, particularly convolutional networks (convnets), onto low-cost hardware. For example, [22] have trained a convnet to recognize 50 classes of environmental sounds on a Sony Spresense MCU.

2. DATASET

The development set for task 3 of BioDCASE 2025 "Bioacoustics for Tiny Hardware" consists of 2 hours and 37 mins worth of audio recordings. See Table 1 for a breakdown.

2.1. Data collection

Yellowhammers are widespread across Eurasia and their songs have attracted the attention of naturalists and scientists for over one and a half centuries (e.g. [23]). They have long been a popular model species, especially for studying dialects in birdsong [24]. They are also an indicator of healthy farmland showing rapid population declines around Europe (e.g., [25]), and their individually specific songs might offer detail noninvasive insights into fast population changes. The yellowhammer songs used for BioDCASE 2025 were initially recorded as a part of sound transmission experiments that investigated how far it is still possible to identify Yellowhammer individuals by their songs. To collect the songs, AudioMoth recorders were placed near the bird's favorite singing posts less than 5 m from the singing birds. The best quality songs were selected and put into one track. The track contained in total 209 songs from 10 individuals, including in total 21 different song types (2–3 song types per male), with c.a. 10 repetitions of each song type.

The track was then played back and re-recorded at 7 different distances ranging from 6.5 to 200 m (see Figure 1 for representative examples of songs at different distances). Taking into account the original recording distance ($<$ 5m), the songs in the closest distance category simulate birds being anything between 6.5m to 11m far away. Note that the signal-to-noise ratio (SNR) is quite low in the last two distance categories, and songs that were not visible on the spectrogram were not included in the dataset, which means that these distance categories may have fewer than 209 samples. The playback was carried out in two different environments (forest and grassland). Altogether, each song is repeated up to 15 times in the dataset (original distance + 7x forest + 7x grassland). These re-recorded songs were annotated and split into separate files and used in the challenge as positive files. In addition, the original recordings of live Yellowhammers were screened, annotated and split to obtain negative files of similar duration. These included songs from other known species or background noise only.

2.2. Data curation

All audio recordings were clipped to 2 seconds and resampled at 16kHz. The dataset was divided into training and validation sets. Yellowhammer recordings were split by individual: 6 for training, 2 for validation, and 2 held out for evaluation. Negative samples were randomly split.

3. CHALLENGE

3.1. Rules

The challenge tasked participants with developing a Yellowhammer bird vocalization detection system for the ESP32-S3-Korvo-2 microcontroller, using the training and validation sets from the Yellowhammer dataset (Section 2), and our baseline framework (Section 3.3). Submissions had to be deployable on the target hardware and capable of real-time audio processing. The submitted models would then be evaluated by organizers on the (hidden) evaluation set using the benchmarking facilities of the baseline framework, on a range of metrics to do with model classification performance as well as its resource efficiency, described in the next section. Rankings

Dataset Component	Recordings	Duration
Training Set		
Total	3,562	1h 58min
Yellowhammer	1,321	44min
Negatives (known species)	956	31min
Negatives (background noise)	1,285	42min
Validation Set		
Total	1,156	37min
Yellowhammer	408	13min
Negatives (known species)	319	10min
Negatives (background noise)	429	14min
Evaluation Set		
Total	1,140	37min
Yellowhammer	394	13min
Negatives (known species)	317	10min
Negatives (background noise)	427	14min

Table 1: Detailed statistics of the BioDCASE 2025 Task 3 dataset. All recordings have 2 second length at 16KHz.

were provided separately for each evaluation metric, encouraging diverse approaches to the performance-efficiency tradeoff inherent in embedded machine learning applications.

3.2. Metrics

Reflecting the real-world tradeoffs of model performance vs computation space and time constraints in embedded devices, the benchmarks measured the following metrics:

- **Average Precision:** A metric that summarizes the precision-recall curve by computing the weighted mean of precisions at each threshold, with the increase in recall from the previous threshold used as the weight. This provides a comprehensive measure of the model's detection accuracy across all confidence levels and object classes.
- **Preprocessing time (ms):** The total execution time of audio preprocessing and feature extraction.
- **Model time (ms):** The total execution time of model inference on the extracted features.
- **Model size (bytes):** The storage footprint of the model weights and architecture.
- **Peak RAM usage (bytes):** The maximum amount of RAM consumed during inference.

3.3. Baseline

To lower the entry barrier and establish fair performance benchmarks, we provided a comprehensive baseline framework through the BioDCASE-Tiny 2025 repository [31]. This framework implements an end-to-end pipeline for developing bird species recognition models in Python and deploying them for benchmarking on an ESP32-S3-Korvo-2 development board. The baseline framework serves as both a starting point for participants unfamiliar with embedded ML development and a reference implementation that demonstrates the expected integration between model development on the host and the deployment and execution of the model on the embedded device.

The system consists of five stages integrated into an automated pipeline (see Figure 2). First, in the data preprocessing stage, the pipeline handles the ingestion and preparation of raw audio files from the Yellowhammer dataset. Second, a feature extraction step transforms the preprocessed audio into log-mel spectrograms for model training, with configurable parameters that participants could adjust to optimize their approaches.

These features were implemented both on the host system and the embedded target in a numerically equivalent manner, ensuring that

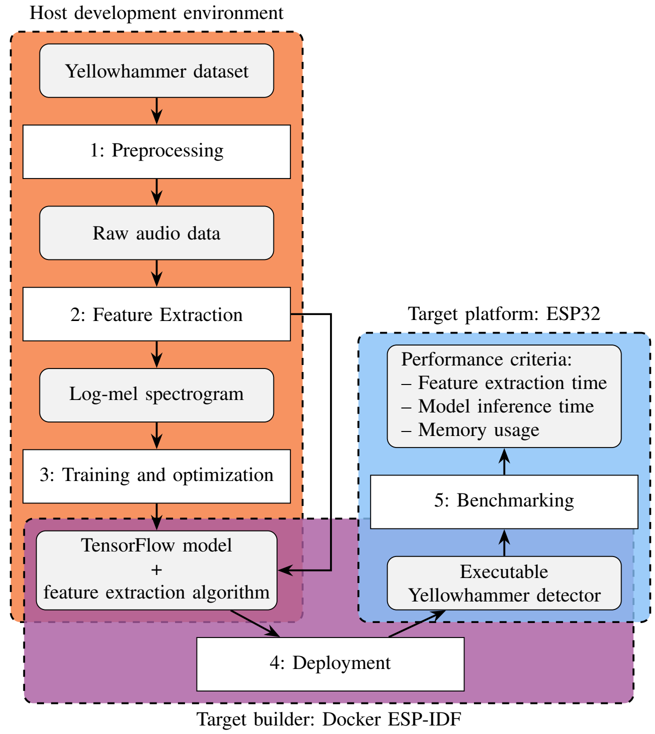

Fig. 2: Flowchart of the baseline development and deployment pipeline for yellowhammer detection, from raw audio input and log-mel spectrogram extraction to model training, deployment, and benchmarking on a constrained embedded device (ESP32).

the deployed model receives identical inputs to those used during training. Third, at the model training stage, participants could define custom architectures as well as a custom training loop if needed. Fourth, at the deployment stage, the pipeline automatically converts the model to TensorFlow Lite format, generates optimized C++ code for the ESP32-S3 platform, and handles firmware compilation and flashing using a Docker-based ESP-IDF tool chain. Finally, at the fifth stage, a benchmarking firmware runs on the embedded target to measure feature extraction time, model inference time, and memory usage, enabling participants to optimize their solutions against the competition's evaluation criteria.

A baseline model was also included as a reference for participants. The model is a lightweight CNN starting with a mobile-net style convolutional and depth-wise convolutional block, followed by a global max pooling layer, dropout and finally a dense layer with softmax activation. The model achieves high average precision on the evaluation data. However, this comes at a rather high cost in terms of model size, RAM usage, and execution time.

Participants in the BioDCASE-Tiny 2025 challenge employed a range of approaches to address the challenge's constraints, and explore different trade-offs to improve on the baseline.

3.4. Neural network architectures

The submitted models demonstrated significant variation in architectural complexity and design philosophy. Three participants built upon MobileNet-inspired architectures for computational efficiency. [26] implemented a slimmed MobileNet V2 with three depthwise separable convolution blocks, while [29] and [30] employed a further stripped-down MobileNet-style CNN.

In contrast, [27] and [28] pursued extreme minimalism through very shallow CNN architectures, with [28] further venturing into SVMs.

	Model	Avg. Precision (%)	Model size (bytes)	Peak RAM usage (bytes)	Preprocessing time (ms)	Model time (ms)	Total execution time (ms)
	Baseline	99.68	14,564	58,360	191.84	302.827	494,667
[26]	Enriched	97.98	16,600	26,844	176.11	24.417	200.527
[26]	Non-enriched	94.49	16,600	26,844	176.22	24.418	200.638
[27]		74.47	7,020	6,744	1.62	16.449	18.067
[28]	Mel-based	97.31	4,192	18,364	34.28	7.387	41.667
[28]	Flux-based	96.52	1,848	780	16.54	0.185	16.725
[29]		98.57	7,776	20,924	73.42	15.164	88.58
[30]	SlimCNN-student	73.68	12,920	30,328	64.84	112.779	177.622

Table 2: Model performance comparison. In the case of multiple submissions from a participant or team, a model identifier is also reported in the Model column. The table reports a range of performance vs resource efficiency tradeoffs explored by the participants, reflecting the range of requirements and constraints in the usage of TinyML for autonomous recording units.

Finally, [32] explored convolutional-recurrent hybrid architectures, demonstrating the potential of temporal modeling for bioacoustic tasks. However, due to the limited amount of operations supported by TensorFlow Lite, their model could not be benchmarked on the embedded target.

3.5. Data augmentation

Several participants recognized the importance of data augmentation for improving model robustness, particularly for low signal-to-noise ratio conditions. [26] compared models trained with and without augmentation, applying pitch shifting (±2 semitones) and white noise addition. [32] implemented a sophisticated dynamic augmentation strategy that adjusted intensity based on SNR estimates, using SpecMixup and Gaussian noise selectively on high-SNR samples to create realistic training scenarios.

3.6. Feature engineering and preprocessing

Participants employed varied feature extraction strategies, with most utilizing log-mel spectrograms but differing significantly in their parameter choices. The majority processed audio to extract mel spectrograms with filter counts ranging from 16 to 64, with window sizes varying from 512 to 2048 samples. Sampling rates and frequency ranges were carefully selected based on domain knowledge, with e.g. [29] focusing on the 3-8 kHz range based on ornithological literature, while [28] resampled the recordings at 14 kHz and targeted specific frequency bands for different models. Notably, [28] also explored alternative features, achieving strong results with spectral flux statistics, while drastically reducing processing time as well as memory footprint.

3.7. Model compression and optimization

Participants employed various strategies to meet the heavy resource constraints. While some focused on architectural simplification and careful hyperparameter selection, limiting compression to post-training model quantization, others implemented sophisticated compression pipelines. [30] demonstrates a comprehensive approach combining knowledge distillation, magnitude-based pruning, and quantization-aware training.

3.8. Computational performance and resource usage

The submitted models achieved a wide range of performance-efficiency tradeoffs: see Table 2. These results collectively demonstrate that effective yellowhammer detection can be achieved across a broad spectrum of computational budgets, with the optimal choice depending on specific deployment constraints and precision/recall requirements.

4. CONCLUSION

Bioacoustics researchers and practitioners have expressed a demand for accurate and energy-efficient algorithms which can operate "on the edge", i.e., on autonomous recording units [18]. From the standpoint of DCASE, meeting this demand requires to rethink the deployment of machine learning systems with hardware constraints in mind. As part of the first edition of the BioDCASE challenge and workshop, we have organized a task of SED on tiny hardware (ESP32-S3) for the yellowhammer *(Emberiza citrinella)*. Certainly, edge computing does not erase all the harms of cloud computing: risks such as rebound effects [33] and human bycatch [34] remain present. Yet, we believe that bioacoustics on tiny hardware is worth exploring in future years. Indeed, as a collective, the participants of BioDCASE 2025 Task 3 have shown that a more ethically responsible practice of computational bioacoustics is technically within reach.

5. ACKNOWLEDGMENT

We thank the organizing committee of BioDCASE, particularly Ben McEwen, Ines Nolasco, and Dan Stowell, for their kind help. We thank Sébastien Faucou, Larissa Sugai, and Renata S. Sousa Lima. PL has been supported by SustainScapes – Center for Sustainable Landscapes under Global Change (grant NNF20OC0059595). YB, IM, and MK have been funded by the European Union under the MSCA-DN Programme, Grant Agreement No. 101116715 (BioacAI). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible.

REFERENCES

[1] J. Kitzes and L. Schricker, "The necessity, promise and challenge of automated biodiversity surveys," *Environmental Conservation*, vol. 46, no. 4, pp. 247–250, 2019.

[2] D. Teixeira, M. Maron, and B. J. van Rensburg, "Bioacoustic monitoring of animal vocal behavior for conservation," *Conservation Science and Practice*, vol. 1, no. 8, p. e72, 2019.

[3] P. Laiolo, "The emerging significance of bioacoustics in animal species conservation," *Biological conservation*, vol. 143, no. 7, pp. 1635–1645, 2010.

[4] S. Madhusudhana, G. Pavan, L. A. Miller, W. L. Gannon, A. Hawkins, C. Erbe, J. A. Hamel, and J. A. Thomas, *Choosing equipment for animal bioacoustic research*. Springer Nature Switzerland AG, 2022.

[5] R. Snyder, M. Clark, L. Salas, W. Schackwitz, D. Leland, T. Stephens, T. Erickson, T. Tuffli, M. Tuffli, and K. Clas, "The soundscapes to landscapes project: Development of a bioacoustics-based monitoring workflow with multiple citizen scientist contributions," *Citizen Science: Theory and Practice*, vol. 7, no. 1, 2022.

[6] M. Marcoux, M. Auger-Methe, E. G. Chmelnitsky, S. H. Ferguson, and M. M. Humphries, "Local passive acoustic monitoring of narwhal presence in the canadian arctic: A pilot project," *Arctic*, pp. 307–316, 2011.

[7] J. Longdon, M. Westerlaken, A. F. Blackwell, J. Gabrys, B. Ossom, A. Ashton-Butt, and E. Acheampong, "Justice-oriented design listening: Participatory ecoacoustics with a ghanaian forest community," in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–12.

[8] D. Teixeira, P. Roe, B. J. van Rensburg, S. Linke, P. G. McDonald, D. Tucker, and S. Fuller, "Effective ecological monitoring using passive acoustic sensors: Recommendations for conservation practitioners," *Conservation Science and Practice*, vol. 6, no. 6, p. e13132, 2024.

[9] R. S. Sousa-Lima, T. F. Norris, J. N. Oswald, and D. P. Fernandes, "A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals," *Aquatic Mammals*, vol. 39, no. 1, pp. 23–53, 2013.

[10] C. M. Wood, J. Champion, C. Brown, W. Brommelsiek, I. Laredo, R. Rogers, and P. Chaopricha, "Challenges and opportunities for bioacoustics in the study of rare species in remote environments," *Conservation Science and Practice*, vol. 5, no. 6, p. e12941, 2023.

[11] M. Vallee, "The science of listening in bioacoustics research: Sensing the animals' sounds," *Theory, Culture & Society*, vol. 35, no. 2, pp. 47–65, 2018.

[12] G. T. Marx, "Ethics for the new surveillance," *The Information Society*, vol. 14, no. 3, pp. 171–185, 1998.

[13] B. Cretois, C. M. Rosten, and S. S. Sethi, "Voice activity detection in eco-acoustic data enables privacy protection and is a proxy for human disturbance," *Methods in Ecology and Evolution*, vol. 13, no. 12, pp. 2865–2874, 2022.

[14] J. E. Parker and S. Dockray, "'All possible sounds': speech, music, and the emergence of machine listening," *Sound Studies*, vol. 9, no. 2, pp. 253–281, 2023.

[15] N. R. Misra, S. Kumar, and A. Jain, "A review on e-waste: Fostering xthe need for green electronics," in *2021 international conference on computing, communication, and intelligent systems (ICCCIS)*. IEEE, 2021, pp. 1032–1036.

[16] M. P. Cenci, T. Scarazzato, D. D. Munchen, P. C. Dartora, H. M. Veit, A. M. Bernardes, and P. R. Dias, "Eco-friendly electronics—a comprehensive review," *Advanced Materials Technologies*, vol. 7, no. 2, p. 2001263, 2022.

[17] V. Lostanlen, A. Bernabeu, J.-L. Béchennec, M. Briday, S. Faucou, and M. Lagrange, "Energy efficiency is not enough: Towards a batteryless internet of sounds," in *Proceedings of the International Audio Mostly Conference*, 2021.

[18] D. Stowell, "Computational bioacoustics with deep learning: a review and roadmap," *PeerJ*, vol. 10, p. e13152, 2022.

[19] L. Schulthess, S. Marty, M. Dirodi, M. D. Rocha, L. Rüttimann, R. H. R. Hahnloser, and M. Magno, "Tinybird-ml: An ultra-low power smart sensor node for bird vocalization analysis and syllable classification," in *2023 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2023.

[20] P. Prince, A. Hill, E. Piña Covarrubias, P. Doncaster, J. L. Snaddon, and A. Rogers, "Deploying acoustic detection algorithms on low-cost, open-source acoustic sensors for environmental monitoring," *Sensors*, vol. 19, no. 3, 2019.

[21] A. P. Hill, P. Prince, E. Piña Covarrubias, C. P. Doncaster, J. L. Snaddon, and A. Rogers, "Audiomoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment," *Methods in Ecology and Evolution*, vol. 9, no. 5, pp. 1199–1211, 2018.

[22] M. Mohaimenuzzaman, C. Bergmeir, I. West, and B. Meyer, "Environmental sound classification on the edge: A pipeline for deep acoustic networks on extremely resource-constrained devices," *Pattern Recognition*, vol. 133, p. 109025, Jan. 2023.

[23] J. J. Oppel, "Ueber den goldammergesang," *Der Zoologische Garten*, vol. 10, pp. 280–286, 1869.

[24] T. Petrusková, L. Diblíková, P. Pipek, E. Frauendorf, P. Procházka, and A. Petrusek, "A review of the distribution of Yellowhammer (Emberiza citrinella) dialects in Europe reveals the lack of a clear macrogeographic pattern," *Journal of Ornithology*, vol. 156, pp. 263–273, 2015.

[25] F. Burns, M. Eaton, D. Balmer, A. Banks, R. Caldow, J. Donelan, A. Douse, C. Duigan, S. Foster, T. Frost, P. Grice, C. Hall, H. Hanmer, S. Harris, I. Johnstone, P. Lindley, N. McCulloch, D. Noble, K. Risely, R. Robinson, and S. Wotton, "The state of the uk's birds 2020," RSPB, BTO, WWT, DAERA, JNCC, NatureScot, NE and NRW, Technical Report, 2020.

[26] S. Espitia Londono, "Biodcase task3: Tinyml models for bird classification," DCASE2025 Challenge, Tech. Rep., June 2025.

[27] C. Walter, "Efficient convolutional neural networks for bioacoustics in tiny hardware," DCASE2025 Challenge, Tech. Rep., June 2025.

[28] I. Oguamanam, L. Machowski, M. Simic, and S. Krishnan, "Targeted feature extraction for bird sound classification systems," DCASE2025 Challenge, Tech. Rep., June 2025.

[29] T. Martin, "Biodcase 2025 challenge technical report," DCASE2025 Challenge, Tech. Rep., June 2025.

[30] D. Naveen, "Convolutional neural network with knowledge distillation for resource-constrained bioacoustics," DCASE2025 Challenge, Tech. Rep., June 2025.

[31] G. Carmantini, F. Förstner, C. Isik, and S. Kahl, "Biodcase-tiny 2025: A framework for bird species recognition on resource-constrained hardware," https://github.com/birdnet-team/BioDCASE-Tiny-2025, 2025.

[32] D. Naveen, "Convolutional gated recurrent network with knowledge distillation for resource-constrained bioacoustics," DCASE2025 Challenge, Tech. Rep., June 2025.

[33] V. Lostanlen and L. Bouchet, "Rebound effects make digital audio unsustainable," in *2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)*. IEEE, 2024, pp. 1–7.

[34] C. Sandbrook, R. Luque-Lora, and W. M. Adams, "Human bycatch: Conservation surveillance and the social implications of camera traps," *Conservation and Society*, vol. 16, no. 4, pp. 493–504, 2018.