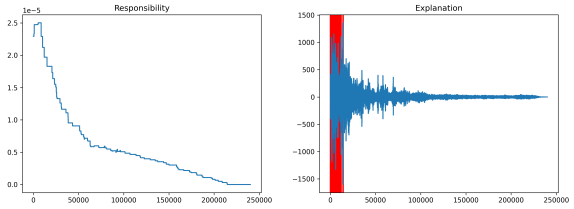
DMRN+20: DIGITAL MUSIC RESEARCH NETWORK
ONE-DAY WORKSHOP 2025
KING'S COLLEGE LONDON
TUE 16 DECEMBER 2025

# Validating Few-shot Bird Vocalisation Detection through Causal Explanations

Nico García-Peguinho[1], David Kelly[2], Melane Navaratnarajah[2],
Hana Chockler[2] and Anna Xambó Sedó[1]
[1]Centre for Digital Music, Queen Mary University of London, UK, n.a.garcia-peguinho@qmul.ac.uk
[2]Department of Informatics/King's College London, UK




(a) Causal responsibility for "blues" (b) Contrastive explanation for "blues"
Figure 1: Removing frequencies in red causes misclassification to "classical" despite the signal retaining blues perceptual qualities, revealing the model's reliance on narrow feature sets.

***Index Terms*—** Bioacoustics, Explainable AI, Few-shot Learning, Causal Explanations, Biodiversity Monitoring

## I. INTRODUCTION

Biodiversity monitoring using deep learning on bioacoustic data is limited by sparse labeled data. Few-shot learning addresses this, but raises the question: *to what extent do these models learn meaningful acoustic features or exploit spurious patterns?* We investigate whether REX, a causal explainability method, can identify which acoustic features few-shot models rely on. We validate this on DCASE 2024 Task 5 bird vocalisations [1] before exploring Sensing the Forest natural recordings from Alice Holt forest, UK.

## II. REX

REX [2] is a tool for explainble AI (XAI) based on actual causality [3]. Unlike popular XAI tools such as LIME [4] or SHAP [5], which produce *saliency maps* measuring feature contribution, REX finds causal explanations by isolating input features required to reproduce the original classification. These isolated features are a *sufficient cause* for that classification. We adapt and apply REX to audio data for the first time. Figure 1 shows a preliminary example of our interpretability method applied to a genre classification model, highlighting the frequencies classified as "blues".

## III. SENSING THE FOREST

Sensing the Forest is a project exploring arts, science, and climate change. Two DIY, solar-powered, off-grid audio recorders [6] were deployed in Alice Holt forest to capture soundscape recordings over a year, aiming to support biodiversity monitoring and climate awareness through community science [7]. This motivates our need for interpretable, validated models before applying bird vocalisation detection to the Sensing the Forest dataset.

## IV. APPROACH

We will train prototypical networks on DCASE 2024 Task 5 BirdVox data using the official baseline system [8]. We will then apply REX to the trained model to identify minimal sufficient spectral-temporal features for bird event detection. Critically, we validate whether identified causal features align with established ornithological acoustic features, ensuring the model learns meaningful rather than spurious patterns.

## V. EXPECTED RESULTS & SIGNIFICANCE

We expect causal feature visualisations and audio examples revealing which frequency-time regions drive bird event detection. This represents the first application of REX to bio-acoustic event detection, providing an interpretable validation framework for few-shot models on bio-acoustic audio data.

## VI. ACKNOWLEDGMENTS

David A. Kelly was supported by CHAI—the EPSRC Hub for Causality in Healthcare AI with Real Data (EP/Y028856/1). Sensing the Forest receives funding from the UKRI Arts and Humanities Research Council (AH/X011585/2).

## VII. REFERENCES

[1] DCASE, "Few-shot Bioacoustic Event Detection - DCASE," https://dcase.community/challenge2024/task-few-shot-bioacoustic-event-detection.
[2] H. Chockler, D. A. Kelly, D. Kroening, and Y. Sun, "Causal explanations for image classifiers," *arXiv preprint arXiv:2411.08875*, 2024.
[3] J. Y. Halpern, *Actual Causality*. The MIT Press, 2019.
[4] M. T. Ribeiro, S. Singh, and C. Guestrin, ""Why should I trust you?" Explaining the predictions of any classifier," in *Knowledge Discovery and Data Mining (KDD)*. ACM, 2016, pp. 1135–1144.
[5] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017, pp. 4765–4774.
[6] L. Marino and A. Xambó Sedó, "Developing DIY Solar-Powered, Off-Grid Audio Streamers for Forest Soundscapes: Progress and Challenges," *CHIME Annual Conference 2024*, 2024.
[7] A. Xambó Sedó, P. Batchelor, L. Marino, G. Roma, M. Bell, and G. Xenakis, "Soundscape-based music and creative AI: Insights and promises," *UKAIRS 2025*, 2025.
[8] J. Liang, I. Nolasco, B. Ghani, H. Phan, E. Benetos, and D. Stowell, "Mind the Domain Gap: A Systematic Analysis on Bioacoustic Sound Event Detection," Mar. 2024.