

ARTICLE TYPE

Identifying Birds through Sound: Challenges and Advances in Machine Learning based Methods

Ali Saudi¹ | Adham Ashton-Butt² | Sebastian Ziegler¹ | Hanna Lukashevich¹ | Jakob Abeßer¹

¹Semantic Music Technologies Group,
Fraunhofer Institute for Digital Media
Technologies (IDMT), Ilmenau, Germany

²British Trust for Ornithology (BTO), The
Nunnery, Thetford, UK

Abstract

Better monitoring of animal populations is vital to understanding, mitigating, and reversing the global biodiversity crisis. Passive Acoustic Monitoring can potentially solve many of the logistical and resource limitations of widespread biological monitoring. However, accurate and scalable machine learning species classifiers are needed to process the vast amounts of data produced. Birds are widely recognised as important indicators of biodiversity, and this article provides a concise overview of cutting-edge deep learning methods for acoustic bird detection. We emphasise the importance and complexity of automating bird call detection and classification and cover key aspects such as data acquisition, signal pre-processing, audio feature representation, and advanced neural network architectures, showcasing the evolution and effectiveness of these techniques for bird recognition. We also highlight current research trends and suggest directions for advancing this rapidly growing field to enhance avian monitoring.

KEYWORDS

Bird sound classification, Bioacoustics, Deep learning, Convolutional neural networks

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Funding acquisition: Hanna Lukashevich, Jakob Abeßer, Adham Ashton-Butt

Conceptualization: Jakob Abeßer, Ali Saudi, and Sebastian Ziegler.

Project administration: Hanna Lukashevich

Supervision: Jakob Abeßer

Writing: Ali Saudi, Jakob Abeßer, Sebastian Ziegler (**original draft**) + Jakob Abeßer, Adham Ashton-Butt, Hanna Lukashevich (**review & editing**).

All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This work was supported by funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101081964.

Abbreviations: CNN, Convolutional neural network; PCEN, Per-Channel energy normalization; HMM, Hidden markov model.