A Review of Audio Classification using Machine Learning: A Systematic Literature Review

Stephan C Kokkas – Deakin University

Tony Zhai – Deakin University

Aishwarya Mariselvam- Deakin University

Project Echo

***Audio Classification using Machine Learning: A Systematic Literature Review***

Audio classification in the scope of our project will be to unobtrusively classify the different types of species in a rainforest. However, one of the benefits of using AI / ML to classify a specific type of data, is that it can easily be transferred to data in a completely different domain – using some fine-tuning techniques. For example, models developed for the research of cardiovascular diseases using audio samples of the heart can be translated and slightly adjusted to fit the domain of audio samples of sound producing animals.

References

[1]

L. Mutanu, J. Gohil, K. Gupta, P. Wagio, and G. Kotonya, “A Review of Automated Bioacoustics and General Acoustics Classification Research,” *Sensors*, vol. 22, no. 21, p. 8361, Oct. 2022, doi: [10.3390/s22218361](https://doi.org/10.3390/s22218361).

[2]

T. Andersson, “Audio classification and content description,” p. 63.

[3]

H. Sinha, V. Awasthi, and P. K. Ajmera, “Audio classification using braided convolutional neural networks,” *IET signal process.*, vol. 14, no. 7, pp. 448–454, Sep. 2020, doi: [10.1049/iet-spr.2019.0381](https://doi.org/10.1049/iet-spr.2019.0381).

[4]

F. J. Bravo Sanchez, M. R. Hossain, N. B. English, and S. T. Moore, “Bioacoustic classification of avian calls from raw sound waveforms with an open-source deep learning architecture,” *Sci Rep*, vol. 11, no. 1, p. 15733, Dec. 2021, doi: [10.1038/s41598-021-95076-6](https://doi.org/10.1038/s41598-021-95076-6).

[5]

W. Chen, Q. Sun, X. Chen, G. Xie, H. Wu, and C. Xu, “Deep Learning Methods for Heart Sounds Classification: A Systematic Review,” *Entropy*, vol. 23, no. 6, p. 667, May 2021, doi: [10.3390/e23060667](https://doi.org/10.3390/e23060667).

[6]

S. Sahoo, P. Dash, B. S. P. Mishra, and S. K. Sabut, “Deep learning-based system to predict cardiac arrhythmia using hybrid features of transform techniques,” *Intelligent Systems with Applications*, vol. 16, p. 200127, Nov. 2022, doi: [10.1016/j.iswa.2022.200127](https://doi.org/10.1016/j.iswa.2022.200127).

[7]

A. Bansal and N. K. Garg, “Environmental Sound Classification: A descriptive review of the literature,” *Intelligent Systems with Applications*, vol. 16, p. 200115, Nov. 2022, doi: [10.1016/j.iswa.2022.200115](https://doi.org/10.1016/j.iswa.2022.200115).

[8]

L. D. Pha, “Robust Deep Learning Frameworks For Acoustic Scene and Respiratory Sound Classification,” Open Science Framework, preprint, Oct. 2021. doi: [10.31219/osf.io/d2tzb](https://doi.org/10.31219/osf.io/d2tzb).

[9]

A. F. R. Nogueira, H. S. Oliveira, J. J. M. Machado, and J. M. R. S. Tavares, “Sound Classification and Processing of Urban Environments: A Systematic Literature Review,” *Sensors*, vol. 22, no. 22, p. 8608, Nov. 2022, doi: [10.3390/s22228608](https://doi.org/10.3390/s22228608).

Recently, technology has developed a lot, especially in the field of Machine Learning (ML), which is useful for reducing human work. In the field of artificial intelligence, ML integrates statistics and computer science to build algorithms that get more efficient when they are subject to relevant data rather than being given specific instructions Machine learning is commonly used in diverse fields to solve difficult problems that cannot be readily solved based on computer approaches. Recently, these advances in machine learning have helped a lot with sound classification, and sound recognition has shown to be a strong value in automating these tasks To say it another way, birds can make two basic sounds: Call and song. While this approach is time consuming, machine learning approaches may also be useful in establishing differentiating between the different species, even after that since it is done on a species of birds that are still not thought to be discernable. However, machine learning's usage of bird classification has only been examined for a small number of species or mannequin processing on the assumption that it can be applied in the real world only through numerical simulation or hand recordings. The results have proven unpractical for ecologists but can be useful for many people of a wide variety of professions

Capturing and automatically recognizing the acoustic emission resulting from typical behavior, i.e., locomotion, feeding, etc., of the target pests. After acquisition the signals are amplified, filtered, parameterized, and classified by advanced machine learning methods on a portable computer. Specifically, we investigate an advanced signal parameterization scheme that relies on variable size signal segmentation. The feature vector computed for each segment of the signal is composed of the dominant harmonic, which carries information about the periodicity of the signal, and the cepstral coefficients, which carry information about the relative distribution of energy among the different spectral sub-bands. This parameterization offers a reliable representation of both the acoustic emissions of the pests of interest and the interferences from the environment.

The authors have declared that no competing interests exist for North Atlantic right whales.