Who's shaking? On using machine learning to detect vibrational signals in laser vibrometry recordings

Marolt, Matija[1*]; Šturm, Rok[2]; Juan José, López Díez[2]; Pesek, Matevž[1]

[1]University of Ljubljana, Faculty for Computer and Information Science, Slovenia
[2]National Institute of Biology, Slovenia
[*]Presenting author, e-mail: matija.marolt@fri.uni-lj.si

The study of vibroscapes involves the analysis of vibroscape recordings, which are gathered in the field for example with laser vibrometry. However, their study is time consuming, as the vibrational signals of different species must first be manually annotated in the recordings.

Our goal is to simplify and speed up annotations by using modern machine learning methods. We analyze vibrational signals of insects living in hay meadows. Altogether, over 100 hours of recordings were analyzed and signals belonging to different species annotated. The vibrational signals are grouped into four distinct vibrational signal types: pulse, train, harmonic and complex, and we develop methods for automatic detection of vibrational signals in the recordings and their classification into the four signal types.

We base our detection and classification models on deep neural networks, which are the preferred choice for pattern recognition tasks in image and audio processing, including bioacoustics. Specifically, we chose to use the Jasper architecture, which is an end-to-end convolutional neural acoustic model. The model was chosen due to its simplicity and a small number of trainable parameters (approx. 300,000 in our models).

The models were trained on a dataset containing four second long vibroscape snippets, where each snippet contained either a vibrational signal or background noise. They were trained to recognize the presence of a vibrational signal in the recording (either any signal, or signal of a particular type). We measured the accuracy of the models on a validation set using standard classification measures of precision, recall and F1. Currently, the best models yield the F1 score of 0.89 for detecting the presence of a vibrational signal, 0.87 for detecting complex, 0.85 for pulse, 0.83 for train and 0.7 for harmonic vibrational signals.