Detecting Pollinators from Stem Vibrations Using a Neural Network

Žan Ambrožič
za44564@student.uni-lj.si
Faculty of Mathematics and Physics, University of Ljubljana
Ljubljana, Slovenia

Lorenzo Bianco
l.bianco@unito.it
Department of Life Science and System Biology, University of Turin
Turin, Italy

Rok Šturm
rok.sturm@nib.si
National Institute for Biology
Ljubljana, Slovenia

David Susič, Maj Smerkol, Anton Gradišek
Department of Intelligent Systems, Jožef Stefan Institute
Ljubljana, Slovenia

Abstract

Passive sensing of pollinator activity is important for biodiversity monitoring and conservation, yet conventional acoustic or visual methods produce large amounts of data and face deployment challenges. In this work, we present initial results on investigating stem vibration as an alternative signal for detecting pollinator presence on flowers. Vibration recordings were collected with a laser vibration instrument from various flower species at multiple locations in Slovenia, totaling approximately 14 hours, of which 3 hours were expert-annotated as positive (insect activity present). The task was formulated as a binary classification problem: determining whether a vibration segment corresponds to a pollinator physically touching the flower. Using a neural network model, performance was evaluated with five-fold cross-validation across three experiments: (i) using a balanced subset, (ii) using the full dataset, and (iii) using the full dataset with heuristic prediction smoothing. On the balanced subset, the model achieved an average F1-score of 0.86 ± 0.06; on the full dataset, 0.62 ± 0.07; and with heuristic smoothing, 0.69 ± 0.11, demonstrating both the feasibility of vibration-based detection and the benefits of post-processing. Beyond binary detection, future work will focus on species- and activity-level classification. Ultimately, the goal is to develop lightweight vibration detectors deployable directly on plants, enabling scalable estimation of pollinator visitation rates in natural and agricultural environments.

Keywords

stem vibrations, pollination, neural networks, buzz detection, spectrograms

1 Introduction

Europe supports a rich diversity of wild pollinators among them 2,051 species of bees and 892 species of hoverflies. Collectively, pollinators provide a wide range of benefits to society including more than €15 billion per year contribution to the market value of European crops, pollinating around 78 percent of wild flowering plants. This pollination service ensures healthy ecosystem functioning and maintains wider biodiversity as well as culturally important flower-rich landscapes [1]. Many reviews highlight the global decline in insects [2], [3] and in particular wild bees [4], [5]. Internationally, the UN Intergovernmental science-policy Platform on Biodiversity and Ecosystem Services (IPBES) and the Convention on Biological Diversity (CBD) highlighted the need to collect long-term high-quality data on pollinators and pollination services in order to direct policy and practice responses to address the decline. There were already some attempts to monitor pollinators' activity from sound/soundscapes recordings (e.g. [6]). Here, we explore for the first time to monitor pollination activity by using vibroscape recordings [7] from flowering plants which are visited by different pollinators. We investigated the possibility of neural networks for automatic detection of pollinator visits on flowers.

2 Dataset

The dataset comprises vibration waveforms from flowers, which were used for model training, and auxiliary audio and camera recordings collected for labeling and species identification. All recordings were obtained during July and August 2024 at various locations in Slovenia. The vibrations were measured using a VibroGo (Polytec, Germany) laser vibration instrument, which has an operational range of up to 30 m and can detect movements up to 6 m s^{-1} at frequencies up to 320 kHz. For this study, measurements were performed at close range, with a frequency span of 0–24 kHz and a sampling rate of 48 kHz.

For the measurements, the flower stem was fixed to a pole to minimize large movements, and a small piece of reflective foil was attached slightly below the flower to enable the laser vibrometer to capture fine vibrations caused by insect activity. Our data acquisition setup is shown in Figure 1.

The dataset comprised vibration recordings of up to 10 minutes each, collected from flowers belonging to the genera *Calystegia*, *Cichorium* (the majority of samples), *Crepis*, *Epilobium*, *Knautia*, *Leontodon*, *Lotus*, *Pastinaca*, and *Trifolium*. In total, the recordings amounted to approximately 14 hours, of which 3 hours were annotated for insect activity (as positive), while the rest did not contain insect activity and was considered negative. Labeling was performed in Raven Pro by expert annotators, who used synchronized audio and video recordings to confirm insect presence and identify species. Each recording was annotated with time intervals indicating insect activity, insect species, activity type, and, when relevant, additional notes. For the purpose of this study, where we are only interested in binary classification of detecting pollinators, all intervals with any insect activity which included physically touching the flower were labeled as 1, and 0 otherwise.

Information Society 2025, Ljubljana, Slovenia

https://doi.org/10.70314/is.2025.skui.5707

Figure 1: Data acquisition setup for recording vibration signals, audio, and visual recordings from flowers.

Labeled intervals were cut into clips of one second with 0.1-second overlap (positive instances), whereas unlabeled portions were similarly divided and treated as negative instances. To balance the dataset, the negative instances were randomly downsampled. Some negative instances contained environmental noise, such as speech, machinery, or wind, and wind noise was occasionally present in positive instances. Examples of vibration signals from honey bee foraging and from wind are shown in Figures 2 and 3, respectively. The final balanced subset consisted of 7334 positive and 8664 negative instances. The positive data distribution by insects is given in Table 1.

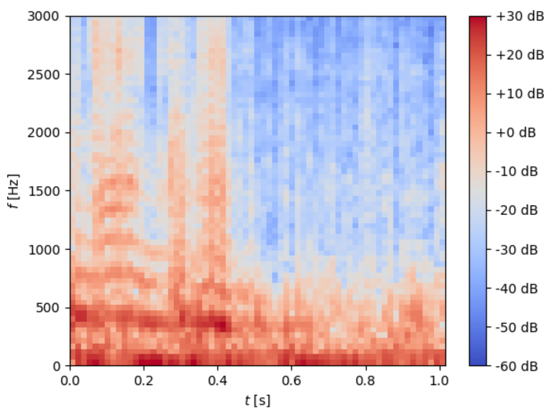

Figure 2: Sample spectrogram of honey bee foraging (positive)

3 Methodology

The objective of this study was to assess whether stem vibrations can be used to detect the presence of pollinators on flowers. From a machine learning perspective, the problem was framed as a binary classification task: distinguishing between the presence and absence of insects in physical contact with the flower. The methodology consisted of initial recoding of waveforms and labeling, preprocessing the data, selecting the appropriate neural network architecture, and training and evaluating the model.

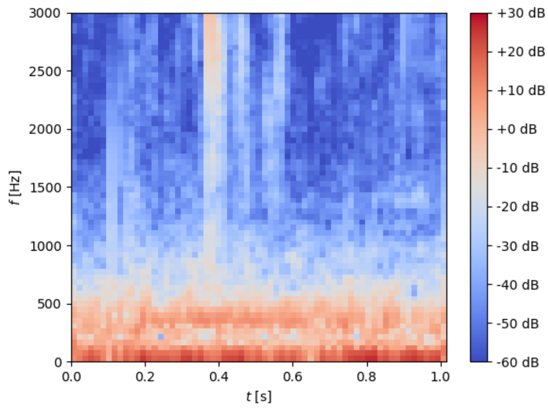

Figure 3: Sample spectrogram of light wind blowing (negative)

Table 1: Number of labels and the corresponding number of instances by insect.

Insect	Number of labels	Instances
fly	76	4146
honey bee	253	1688
wild bee	98	1307
hoverfly	82	155
bumble bee	14	24
wasp	3	9
moth	1	5
Total	527	7334

3.1 Data Preprocessing

First, the instances that were shorter than one second (in cases where the expert-labeled interval was shorter than one second) were padded. After that, all instances were then converted into Mel spectrograms of size 64x64 using fast Fourier transform with frequency range 0–3 kHz.

3.2 Model Architecture

For the model, a network of residual blocks in combination with convolution was used. It is a smaller version of some ResNet (e.g. ResNet 18) models. Residual blocks offer efficient reuse of features across the layers. As shown in Figure 4, the input spectrogram goes through a 3x3 convolution, followed by three residual blocks, before final pooling. The residual block, shown in Figure 5, consists of two 3x3 convolutions to identify features and residual path only uses stride to match the shape before addition.

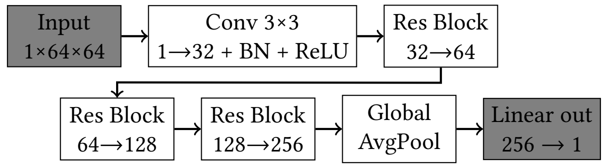

Figure 4: Model Architecture

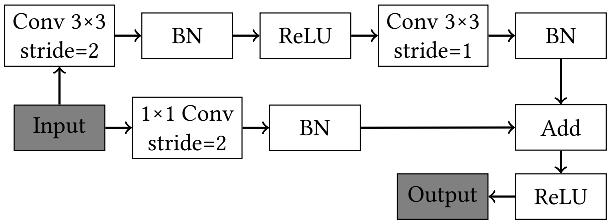

Figure 5: Residual Block (Res Block) in Figure 4

To prevent overfitting and to enable extended training, dropout of 0.5 was used, which improved performance more than data augmentation (and was also computationally more efficient).

3.3 Model Training Settings

The model was trained by using the binary cross-entropy loss. Optimization was performed with Adam optimizer with learning rate 10^{-4} and weight decay 10^{-5}. A batch size of 16 and an epoch number of 30 were used.

4 Evaluation Metrics

We used standard performance evaluation metrics: accuracy, precision, recall and F1-score, which were computed from the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

In confusion matrices, we used relative numbers samples for colors instead of absolute (which are only listed), because there was much more negatives than positives in detection test. Relative shares are based on true labels (e.g. fraction of FN among all negatively labeled).

4.1 Experiments

The model was evaluated in three experimental settings, all using 5-fold cross-validation. Instances originating from the same recording were always assigned to the same fold to better reflect real-world variability. Training and testing were repeated five times, each with a different fold held out for testing and the remaining folds used for training. Reported results are averages across the five folds.

4.1.1 Balanced labeled subset. In the first experiment, called "Subset", only the manually labeled subset of the dataset was used. This consisted of the 7334 positive and 8664 negative instances as described above. These were treated as balanced binary classification data and evaluated directly.

4.1.2 Full dataset with raw labeling. In the second experiment, called "Full data (raw)", the entire dataset was included by segmenting recordings into 1.0 s windows with a step size of 0.1 s. Expert annotations were then used to assign labels to these windows, yielding a much larger evaluation set. However, such raw labeling frequently introduced short, isolated positive or negative events that were likely erroneous. When the model predicted such isolated events, performance metrics were underestimated, as the evaluation framework treated them as genuine labels. This motivated the introduction of a heuristic smoothing procedure.

4.1.3 Full dataset with heuristic labeling. The third experiment, called "Full data (heuristics)", used the same sliding-window segmentation as raw labeling experiment, but applied a heuristic smoothing procedure to adjust labels. The aim was to reduce the influence of short, likely erroneous events while preserving longer, fragmented signals as single detections. Two rules were applied:

- If the model predicted at least 10 consecutive positive windows (equivalent to 1.0 s), the entire interval was relabeled as positive.
- If at least 82% of 50 consecutive windows (equivalent to 5.0 s) were predicted as positive, the entire interval was relabeled as positive.

These empirically determined thresholds suppressed short false positives while ensuring that extended pollinator events with intermittent weak signals were still detected as continuous segments. Finally, because the sliding window (1.0 s) exceeded the step size (0.1 s), prediction timestamps were shifted backward by 0.5 s to align the window centers with the expert annotations.

5 Results and Discussion

The results of all three experiments are shown in Table 2 along with the confusion matrices in Figure 6.

Table 2: Results of all experiments. The numbers represent the average ± standard deviation across five folds in the cross-validation.

	Subset	Full data (raw)	Full data (heur.)
Accuracy	0.87 ± 0.03	0.80 ± 0.02	0.86 ± 0.05
Precision	0.85 ± 0.09	0.54 ± 0.11	0.68 ± 0.15
Recall	0.87 ± 0.04	0.75 ± 0.11	0.73 ± 0.13
F1-score	0.86 ± 0.06	0.62 ± 0.07	0.69 ± 0.11

The results show that there was a significant reduction in performance when we switched from using the balanced subset to recordings from the full dataset. There are several possible sources of error: labels are annotated on waveform and samples are extracted in the way that the whole non-padded (therefore non-silent) part is either positive either negative, furthermore, prediction for a specific time t is generated based on the window, beginning at t and ending at $t + 1$ s, which might lead to inaccuracies at edges of labels although we shifted the time to match it as good as possible. There are also no other insects or activities in samples, which occur in full recordings and are

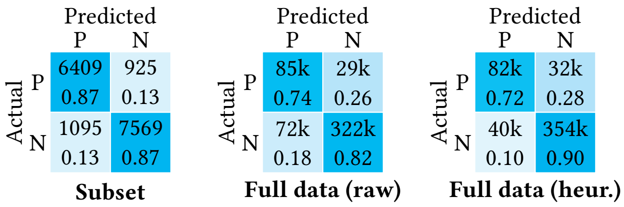

Figure 6: Confusion matrices of all 3 experiments, described in section 4.1

sometimes falsely positive. It is important to note that the "Full data" is not a balanced set (while "Subset" is) and is meant as a test for a real-world scenario, where conditions and frequency of pollinators with them vary on short time scales (hours), which makes loss balancing (which would reduce the gap between recall and precision) in practice very difficult. For this reason, we did not use it and we left the thresholds the same as in the "Subset" experiment, so the results serve as a valid estimation of the performance in reality. Figure 7 shows how heuristics helped the model by smoothing out the short erroneous predictions, resulting in improved performance. To improve model performance even further, additional heuristic filters may be added.

6 Conclusion

We presented initial results on the feasibility of detecting pollinator presence on flowers from stem vibration recordings using machine learning methods. We evaluated models under three experimental settings: a balanced labeled subset, the full dataset with raw expert annotations, and the full dataset with heuristic label smoothing. The results demonstrate that pollinator activity can be reliably inferred from vibration signals, with heuristic post-processing substantially reducing the impact of isolated erroneous predictions and improving the robustness of detection.

Future work will focus on extending the models beyond binary detection towards classification of pollinator species and potentially of behavioral activities. From an applied perspective, the long-term goal is to develop lightweight vibration detectors that can be mounted directly on plants to automatically register pollinator visits. Deploying a small number of such sensors in a field or meadow would enable scalable estimation of pollinator abundance and activity, providing a valuable tool for biodiversity monitoring and conservation studies.

Acknowledgements

The authors acknowledge the funding from the Slovenian Research and Innovation Agency, Grants PR-10495, P1-0255, J7-50040, Z1-50018, the Basic core funding P2-0209, and the support received from the Erasmus+ Traineeship programme.

References

[1] European Commission. Joint Research Centre. 2021. *Proposal for an EU pollinator monitoring scheme.* Publications Office, LU. DOI: 10.2760/881843.

[2] David L. Wagner. 2020. Insect declines in the anthropocene. *Annual Review of Entomology*, 65, 1, (Jan. 2020), 457–480. DOI: 10.1146/annurev-ento-011019-025151.

[3] Roel van Klink, Diana E. Bowler, Konstantin B. Gongalsky, Ann B. Swengel, Alessandro Gentile, and Jonathan M. Chase. 2020. Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science*, 368, 6489, (Apr. 2020), 417–420. DOI: 10.1126/science.aax9931.

[4] J. C. Biesmeijer et al. 2006. Parallel declines in pollinators and insect-pollinated plants in britain and the netherlands. *Science*, 313, 5785, (July 2006), 351–354. DOI: 10.1126/science.1127863.

[5] Luísa Gigante Carvalheiro et al. 2013. Species richness declines and biotic homogenisation have slowed down for <scp>nw</scp>-european pollinators and plants. *Ecology Letters*, 16, 7, (May 2013), 870–878. Yvonne Buckley, editor. DOI: 10.1111/ele.12121.

[6] David Susič, Johanna A. Robinson, Danilo Bevk, and Anton Gradišek. 2025. Acoustic monitoring of solitary bee activity at nesting boxes. *Ecological Solutions and Evidence*, 6, 3, e70080. e70080 ESO-24-09-164.R1. eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1002/2688-8319.70080. DOI: https://doi.org/10.1002/2688-8319.70080.

[7] Rok Šturm, Juan José López Díez, Jernej Polajnar, Jérôme Sueur, and Meta Virant-Doberlet. 2022. Is it time for ecotremology? *Frontiers in Ecology and Evolution*, 10, (Mar. 2022). DOI: 10.3389/fevo.2022.828503.

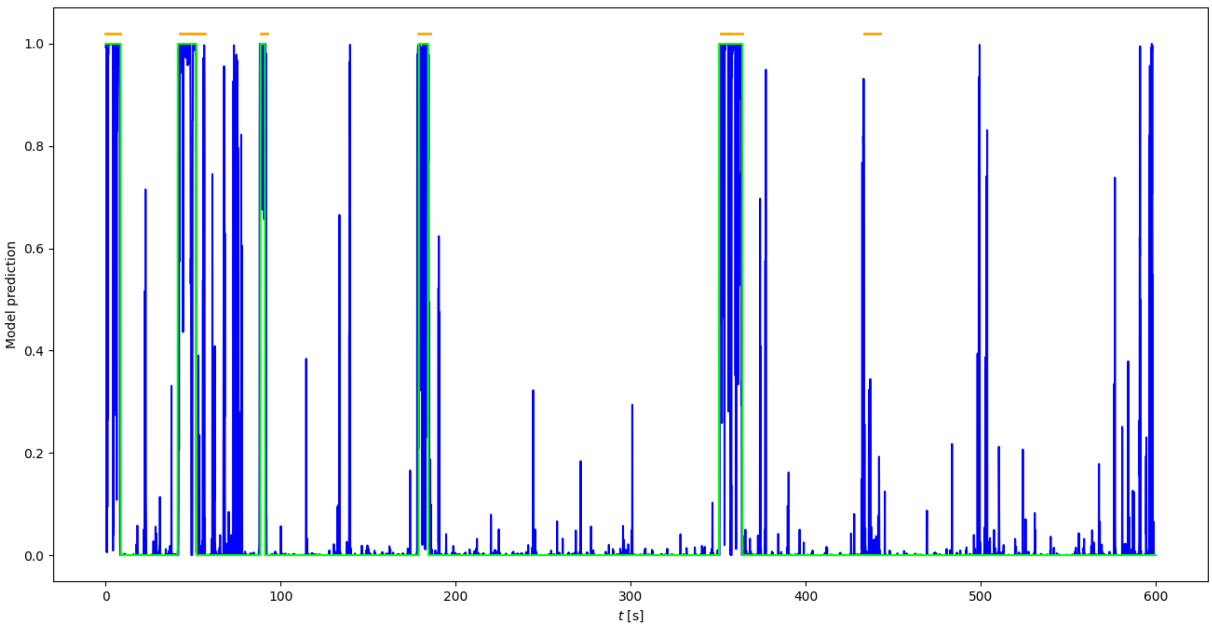

Figure 7: Output example: (blue) model prediction, (green) heuristic filter, (yellow) expert labels.